Journal of Communication, 2022, **00**, 1–12
https://doi.org/10.1093/joc/jqac047
Original Article

International Communication Association
OXFORD

Modeling news recommender systems' conditional effects on selective exposure: evidence from two online experiments

Erik Knudsen [1]*

[1]MediaFutures: Research Centre for Responsible Media Technology and Department of Information Science and Media Studies, University of Bergen, Bergen, Norway

*Corresponding author: Erik Knudsen. Email: erik.knudsen@uib.no

Abstract

Under which conditions do news recommender systems (NRSs) amplify or reduce selective exposure? I provide the Recommender Influenced Selective Exposure framework, which aims to enable researchers to model and study the conditional effects of NRSs on selective exposure. I empirically test this framework by studying user behavior on a news site where the choice environment is designed to systematically influence selective exposure. Through two preregistered online experiments that simulate different NRSs and unobtrusively log user behavior, I contribute empirical evidence that an NRS can increase or decrease the chance that selective exposure occurs, depending on what the NRS is designed to achieve. These insights have implications for ongoing scholarly debates on the democratic impact of NRSs.

Keywords: selective exposure, news recommender systems, filter bubble, online experiment, news nudging

Concerns over the possible adverse effects of artificial intelligence and news recommender systems (NRSs) on democracy have spawned a raft of literature on the extent to which online news environments are characterized by filter bubbles in which users mainly encounter information that conforms to their own political attitudes and beliefs. Decades of research have repeatedly shown that individuals prefer political news that supports their views, known as *selective exposure* (Knobloch-Westerwick, 2015; Lodge & Taber, 2013). While selective exposure can occur both offline and online (Knobloch-Westerwick, 2015), some fear that NRSs that filter the abundance of information online will enable selective exposure on an unprecedented scale by learning audiences' preferences for like-minded content and causally increase the chance that selective exposure occurs by promoting news that audiences agree with at the expense of the news they disagree with (Dylko, 2016; Pariser, 2011).

At least three perspectives with three different expected outcomes regarding the causal effects of NRSs on selective exposure can be distinguished in the prior literature. The first perspective argues that NRSs lead to increased selective exposure (e.g., Dylko, 2016; Pariser, 2011). The second perspective argues that NRSs do not increase selective exposure among the general public (e.g., Bruns, 2019b; Dahlgren, 2021; Zuiderveen Borgesius et al., 2016). The third perspective argues that the democratic consequences of NRSs will likely depend on what the NRS are programmed to achieve (e.g., Helberger, 2019) and that NRSs can be programmed to *decrease* selective exposure by promoting viewpoint diversity through nudging (e.g., Garrett & Resnick, 2011; Mattis et al., 2022). While these three perspectives indicate that the direction and presence of the causal effects of NRSs are conditional on certain factors, they have thus far largely been studied as independent strands of the literature. Missing from this literature is a theoretical understanding of the precise conditions under which NRSs are expected to causally increase or decrease the chance that selective exposure occurs.

In this article, I contribute a framework that aims to enable researchers to model and study the conditional effects of NRSs on selective exposure, incorporating the three expected outcomes described above. The framework builds on the theoretical assumption that, while structure and agency are mutually constituted, structural factors can, under certain conditions, influence how audiences are likely to behave (Webster, 2014). More specifically, it builds on the theoretical assumption that NRSs can influence users' online behavior depending on variations in what NRSs are designed to achieve (e.g., Garrett & Resnick, 2011; Helberger, 2019; Mattis et al., 2022). I introduce the Recommender Influenced Selective Exposure (RISE) framework, which distinguishes how variations in NRS design decisions are expected to causally increase or decrease the chance that selective exposure will occur. I empirically illustrate the RISE framework through two online experiments with simulated NRSs designed to influence selective exposure by increasing the salience of an article and highlighting information that is likely to be relevant for the reader.

The state of the art of NRSs' causal effects on selective exposure

NRSs are increasingly used on online news sites to make editorial decisions (Diakopoulos, 2019; Mitova et al., 2022). Machine learning algorithms can personalize user experiences by deciding on the selection and order of stories on the front page and by deciding on additional stories to suggest to the reader of a story. The literature separates between user-driven (i.e., users can customize the recommender) and system-driven

Received: 11 February 2022. **Revised:** 4 December 2022. **Accepted:** 5 December 2022

(i.e., software code that does all the customization for the user) customization (Dylko, 2016; Zuiderveen Borgesius et al., 2016). This article focuses on system-driven customizability.

In general, NRSs are an important feature to help online news readers choose among an overwhelming number of available offers and highlight information that users find relevant (Lazer, 2015). NRSs often base their information recommendations about users' preferences according to these users' previous choices or ratings (Karimi et al., 2018). A common class of systems bases its recommendation on the assumption that similar users have similar preferences, known as collaborative filtering, that is, "other people who liked that item, also liked this item." Another key approach is content-based filtering, which bases the recommendations on the type of content the user has preferred in the past. Most NRSs use a hybrid approach, combining content-based and collaborative filtering recommendations (Karimi et al., 2018). These approaches are likely to actively promote articles that resemble those a user has read before. According to Pariser (2011) and Dylko (2016), this has implications for selective exposure. Following their argument, if users are likely to prefer news that caters to their political preferences, this should also increase the chance that selective exposure will surge over time (Dylko, 2016; Pariser, 2011). As NRSs learn from the users' choices, and users make those choices predominantly from the options promoted by the algorithms, a self-reinforcing feedback loop might gradually constrain users' choice to an increasingly narrow and homogeneous selection of news content. Empirical studies of such consequences have focused on whether individuals prefer like-minded information on different platforms, sites, and services (e.g., Bakshy et al., 2015; Beam, 2014; Bechmann & Nielbo, 2018; Haim et al., 2018; Huszár et al., 2022) with mixed results (Bruns, 2019b; Terren & Borge-Bravo, 2021). Among this mixed evidence, at least three different perspectives with different expected outcomes on the causal effects of NRSs on selective exposure can be distinguished.

First, some argue that NRSs increase selective exposure (e.g., Dylko, 2016; Pariser, 2011). For instance, Levy (2021) found, through a randomized controlled trial (RCT), causal evidence that even when users were willing to like the Facebook pages of outlets with which they disagreed, Facebook's algorithm was less likely to show them content from those outlets. Dylko et al. (2017, 2018) found, using an experiment with a convenience sample in which the experimenters unobtrusively captured user behavior on an online mock news site, that NRSs increased selective exposure. Bakshy et al. (2015) found, through analyses of a comprehensive Facebook dataset with self-reported political ideology, that Facebook's NRS decreased exposure to cross-cutting content in the news feed by 15% but that individuals' choices played a more pronounced role in limiting exposure to cross-cutting content than NRSs' ranking of content. A systematic review of the published literature (between 2011 and 2020) further indicated that most prior studies found evidence that selective exposure *occurs* in a range of online news environments that use NRSs (Terren & Borge-Bravo, 2021).

Second, some argue that while selective exposure occurs in online services that use NRSs, an NRS does not causally increase selective exposure (e.g., Bruns, 2019b; Dahlgren, 2021; Zuiderveen Borgesius et al., 2016). A recent literature review concludes that "the forms of algorithmic selection offered by search engines, social media, and other digital platforms generally lead to slightly more diverse news use—the opposite of what the 'filter bubble' hypothesis posits" (Arguedas et al., 2022, p. 4). For instance, Fletcher and Nielsen (2017) found, through a comparative survey on cross-media and cross-platform self-reported news use in six countries, correlational evidence that online news use (including but not limited to platforms and sites using NRSs) has not separated a great share of the audience into bubbles where they do not meet counter attitudinal content. Fletcher et al. (2021) found, through analyses of the tracking data of browsing on desktops or laptops, correlational evidence that biases in news repertoires are primarily a result of self-selection and not of system-driven NRSs; however, they did not study a specific NRS. Moreover, Möller et al. (2018) conducted a simulation experiment that compared the output of different algorithmic news recommendations for the same articles from the same news source. While they did not study viewpoint diversity, they found that several different commonly used features of NRSs tend to lead to a diverse set of recommendations that are on par with human editors.

A third perspective adds a causal expectation to the second perspective and argues that NRSs can causally *reduce* selective exposure (e.g., Bozdag & van den Hoven, 2015; Garrett & Resnick, 2011; Helberger et al., 2018; Mattis et al., 2022). For instance, Beam (2014) found, through a mock gubernatorial election experiment with a convenience sample, that system-driven NRSs decreased selective exposure. By developing four different NRS designs and testing them in an online experiment, Joris et al. (2022) tested the causal effects of four NRS design decisions on diverse news exposures. Similarly, Heitz et al. (2022) developed and tested a smartphone app to test the causal effects of chronological, accuracy-, or diversity-optimized NRSs on user behavior. Although Joris et al. (2022) and Heitz et al. (2022) empirically illustrated that NRSs can be designed to increase exposure diversity, neither of the two studies explicitly studied the causal effects on selective exposure.

Contrary to these three perspectives, this article argues that the question to ask is not necessarily *whether* NRSs causally increase or decrease selective exposure; but rather, *What are the conditions under which NRSs causally increase or decrease the chance that selective exposure occurs?* This question connects the three perspectives, as it allows for all the three perspectives to be true under certain conditions and to be studied simultaneously under a common framework. Despite prior efforts to understand whether NRSs influence selective exposure, the precise conditions under which NRSs amplify or reduce selective exposure remain a puzzle. Prior studies have typically faced an obstacle in terms of their ability to conceptualize an NRS as a causal factor, as researchers will rarely have sufficient access to (Bruns, 2019a) and experimental control over the NRSs they study to determine the conditions under which it influences selective exposure through causal counterfactual interventions (Pearl, 2009). After all, to address the puzzle of the conditional effects of NRSs, analysts should ideally be able to model counterfactual scenarios—that is, scenarios that have not (yet) been realized but could or might under differing conditions. If an NRS on a news site or social media platform influence selective exposure, this effect is likely to arise as a biproduct, as NRSs are rarely developed for the purpose of influencing selective exposure. One such biproduct could be engagement. Levy (2021) found indications of people decreasing their Facebook usage after a nudge to subscribe to counter-attitudinal outlets. If engagement is correlated with selective exposure, Facebook may inadvertently have an incentive to filter counter-attitudinal news if it aims to maximize engagement. Yet, without sufficient access to or experimental control over the NRSs under study, analysts simply do not know the conditions under which the NRSs have the capacity to influence selective exposure. Research that, for instance, aims to test the effects of Google's NRS on selective exposure will rarely be able to disentangle *why* the NRS did or did not influence selective exposure. As different platforms also regularly update their services, the study of NRSs on social media platforms and news services is a study of "black box" technology (Diakopoulos, 2019), which is a moving target that is constantly changing. For instance, Huszár et al. (2022), who conducted an RCT using Twitter's ranking algorithm, noted that the treatment (i.e., the ranking experience of the ranking algorithm) did not remain the same during the data collection period. Thus, without experimental control over NRSs, researchers will have limited opportunities to compare the effects of NRSs, as they will not necessarily know whether the identified effects (or lack thereof) of a certain NRS were due to design changes in the NRS and to what extent their results will hold in future iterations of the NRS under study.

The RISE framework

In this article, I propose a new conceptual framework for modeling the conditional causal effects of NRSs that aims to help researchers bypass the issues discussed above on conceptualizing NRSs as a causal variable: the RISE framework. The premise of this framework is a simple but unexplored proposition: *that the extent to which an NRS increases or decreases the chance that selective exposure occurs is conditional on what the NRS is designed to achieve.* More specifically, I ask the counterfactual causal question: What are the conditions under which NRSs amplify or reduce selective exposure in online news environments, given that they are designed to do so? The framework adds the ability to model the conditional causal effects of NRSs on selective exposure by explicitly focusing on experimentally controlling the NRS design and programming decisions (i.e., what the NRS is designed to promote and how). It builds on and extends two key theoretical arguments from prior literature: (a) that NRSs can causally influence user behavior depending on what the NRSs are designed to achieve (e.g., Garrett & Resnick, 2011; Helberger, 2019; Mattis et al., 2022) and (b) that NRSs can be designed to influence selective exposure through promoting factors that prior work has identified as moderators of selective exposure (Hart et al., 2009). While prior theorizing has argued that an NRS can influence exposure diversity depending on what the NRS is designed to achieve (e.g., Garrett & Resnick, 2011; Helberger, 2019; Mattis et al., 2022), we lack a solid theoretical framework for understanding how different design decisions can condition the *direction and presence* of NRSs' causal effects on *selective exposure*. The advantage of such counterfactual theorizing is that it allows researchers to model NRS design decisions as causal variables that potentially influence selective exposure under certain conditions. Analysts can then focus on testing the effects of different NRS design decisions and, for instance, compare the effects of these design decisions to the effects of the different algorithms currently being used in online news environments to come closer to solving the puzzle of NRSs' conditional effects on selective exposure.

The starting point of the RISE framework—illustrated in Figure 1—is that an actor (e.g., from the industry or the academic community) decides to design an NRS with the goal of causally increasing or decreasing the chance that selective exposure occurs. Such decisions are labeled the "input side" of the framework. After such an NRS has been developed, the "output side" describes three distinctive causal outcomes: (a) selective exposure is expected to occur in certain online environments with and without an NRS. Compared to such environments, the chance that selective exposure occurs is expected to (b) *increase* or (c) *decrease*, depending on whether the NRS is designed to promote attitude-*consistent* or *inconsistent* content, respectively, through the nudging of moderators of selective exposure.

The RISE framework has four key assumptions. First, individuals are likely to prefer attitude-consistent content over attitude-inconsistent content (i.e., selective exposure), but this preference can be moderated by other factors. This assumption adopts the model of dual-process modes of thinking, which distinguishes between how decisions are processed: spontaneous, fast, and effortless (System 1) and slow, deliberative, and effortful (System 2) (Lodge & Taber, 2013). When individuals browse and process news online via System 1, they often rely on certain factors, such as whether they agree or disagree with a news headline, and tend to exert a *confirmation bias* and accordingly prefer information that is consistent with their existing attitudes (Hart et al., 2009; Lodge & Taber, 2013). A growing body of literature explores how selective exposure competes against and is moderated by other factors that influence people's media selection decisions. Several moderators of selective exposure build on Atkin's (1973) idea that information selection is driven by how useful one perceives the information to be. Atkin (1973, as cited in Knobloch et al., 2003) noted that "the individual desires to formulate precise cognitive orientations toward those stimuli that potentially or currently impinge on his well-being" (p. 94). Building on Atkin's (1973) work, a large and growing body of literature on moderators of selective exposure has since found evidence that individuals select news content that is more likely to provide them with "instrumental utility." This is also reflected in a meta-analysis by Hart et al. (2009), which found that while people generally prefer consistent over inconsistent information, inconsistent information can be more, or equally, preferred when it also has a high utility.

The second assumption of the RISE framework builds on the growing literature on conditional selective exposure effects and assumes that an NRS could implement contextual factors that prior studies have shown to moderate selective exposure (i.e., the input side in Figure 1). It assumes that, during the design phase, an NRS can be implemented with moderators of selective exposure and different approaches to "nudging" (Mattis et al., 2022; Thaler & Sunstein, 2008) content to make subtle changes in the choice environment.

The third assumption is that the design decisions on the input side of the RISE framework are likely to influence the output side of the RISE framework (the right side of Figure 1). Depending on the design choices on the input side, news sites (or other online environments) could, in a counterfactual scenario, use system-driven NRSs that are designed to seamlessly either causally decrease or increase selective exposure. For instance, Garrett and Resnick (2011) theorized that technology

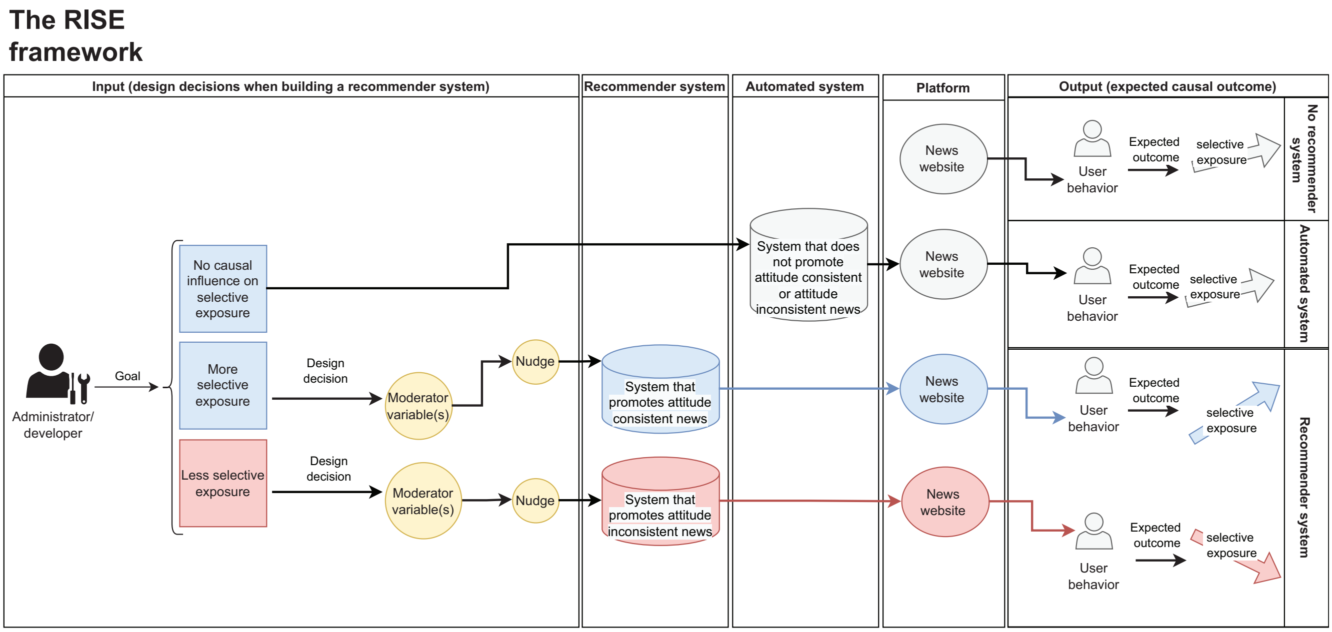

Figure 1. The RISE framework.

equipped with knowledge from prior research on moderators of selective exposure “can nudge individuals slightly in the direction of exposure to challenging viewpoints” (p. 109). More specifically, the RISE framework assumes that NRSs that are programmed to increase (decrease) selective exposure can nudge users to asymmetrically expose themselves to more attitude-consistent (inconsistent) content than attitude-inconsistent (consistent) content.

The fourth assumption is that one can use certain online choice environments as baselines. When a choice environment has a limit in terms of browsing time and number of available articles, and contains attitude-consistent and inconsistent stories but does not systematically promote or demote any of these stories, users are more likely to select attitude-consistent content than attitude-inconsistent content (i.e., engage in selective exposure) (Knobloch-Westerwick, 2015). For analysts to be able to evaluate whether an NRS causally increases or decreases the chance that selective exposure occurs, they need to be able to compare the effects of that NRS to a baseline. As a baseline, the RISE framework recommends using an online choice environment that, on the one hand, likely provides the opportunity for selective exposure to occur by including at least one attitude-consistent article and at least one attitude-inconsistent article for users to select (but not necessarily an identical number of each) and that, on the other hand, is not likely to systematically increase or decrease the chance that selective exposure will occur. One could, for instance, use a news site without an NRS, or an NRS that is designed to randomly distribute news (such as the “Random diversity-based” NRS tested in Joris et al., 2022), as a baseline. Analysts can then test whether selective exposure is more or less likely, given the causal intervention of NRSs.

To model NRSs’ conditional effects on selective exposure through the RISE framework, one should precisely define what to estimate by establishing an estimand that can link the theoretical framework to empirical evidence (Lundberg et al., 2021). In the empirical studies in this article, the estimand is the difference in whether an attitude-consistent news article would be selected over an attitude-inconsistent news article by Internet users if it were:

1) neither made more or less appealing and easy to select than an attitude-inconsistent news article vs.
2) systematically promoted to be more appealing and easier to select than an attitude-inconsistent news article through an NRS vs.
3) systematically demoted to be less appealing and less easy to select than an attitude-inconsistent news article through an NRS.

As I will show in the following section, this estimand has implications for whether the chance that selective exposure occurs is expected to increase or decrease due to the causal intervention of varying NRS design decisions.

Testing NRSs’ conditional effects on selective exposure

Establishing a baseline

The starting point for testing the effects described on the output side of the RISE framework is to establish a baseline. There are at least two baseline scenarios that are relevant for testing the RISE framework. First, I assume that selective exposure is likely to occur when an NRS is absent, and the choice environment is shaped by human editors’ gatekeeping decisions. For such choice environments not to causally influence selective exposure, I assume, building on empirical evidence, that journalistic professionals generally do *not* tend to exert a confirmation bias in terms of what they deem newsworthy (Hassell et al., 2020). Therefore, I assume that human editors are, in general, not likely to promote news stories with which they agree over news stories with which they disagree—at least not due to the attitude consistency of the story. Note, however, that this assumption does not necessarily hold for news outlets with political leanings, as such leanings may lead to an overall bias in the selection or framing of news stories. In this study, I assume, for simplicity, that the news site does not have political leaning. A second possible baseline scenario that should not causally influence selective exposure is a choice environment with an automated system that is designed to *not* make attitude-consistent news or attitude-inconsistent news more or less appealing and easy to select. For instance, in a choice environment where the opportunity to select attitude-consistent or attitude-inconsistent news is equal due to randomization. Two baseline hypotheses follow.

> H1a: When the order of news stories on news sites is determined by human editors, users are more likely to select an attitude-consistent article than an attitude-inconsistent article.
> H1b: When the order of news stories on news sites is random, users are more likely to select an attitude-consistent article than an attitude-inconsistent article.

Causal interventions: promoting factors that increase or decrease selective exposure

The RISE framework assumes that an NRS that implements moderators of selective exposure will be able to increase or decrease the chance that selective exposure occurs depending on what the NRS is programmed to achieve. To test this proposition, I outline how the chance of selecting attitude-consistent news over attitude-inconsistent news is expected to change depending on small but systematic changes in the choice environment that should make attitude-consistent or attitude-inconsistent news more or less appealing to select.

To test how selective exposure can be influenced by an NRS that nudges a known moderator of selective exposure, I highlight the moderating effects of *relevance* and *salience*. Users are more likely to select and read news they find relevant. For instance, the literature on "issue publics" examines individuals who are passionately interested in certain topics and therefore consume more news on them (Feldman et al., 2013). Knobloch-Westerwick et al. (2005) found a positive effect of geographical relevance on headline selection and reading time. Mummolo (2016) found that, building on Atkins' (1973) work on utility, topic relevance reduced selective exposure to congenial sources, indicating that selective exposure (at least to sources) is conditional on relevance. Users are also more likely to select news that is more salient than others through factors such as ordering (i.e., ordered at the top of the page) and picture size. For instance, Loecherbach et al. (2021) found a strong positioning effect, where articles ordered on the upper left of a 3×3 grid were more likely to be selected than articles on the lower right. Building on nudging theory (Thaler & Sunstein, 2008), I assume that one can design an NRS to nudge (a) the relevance moderator variable through, for instance, highlighting whether a given article mentions a place that is geographically close to the user and (b) the salience moderator variable through systematically varying the order, placement, and size (e.g., picture or font size) of an article.

Identifying and promoting articles that are likely to be relevant for users, as well as re-ordering articles, are key parts of the existing NRS technology. If one designs an NRS to influence selective exposure through the RISE framework, relevance recommendations and salience recommendations are highly plausible factors to use to achieve that. Based on the evidence on the effects of salience and relevance on selective exposure reviewed above, it is reasonable to expect that news environments that feature tailor-made system-driven NRSs to promote or demote attitude-consistent or attitude-inconsistent news through relevance recommendations and salience recommendations will influence selective exposure. Note that a decrease in selective exposure and an increase in selective exposure are two different effects that are potentially influenced to a different extent by the relevance recommendation and the salience recommendation. Two separate treatment hypotheses follow.

> H2: Users will be more likely to engage in selective exposure than the two baseline conditions when the attitude-consistent story is relevant and/or salient and the attitude-inconsistent story is not.
> H3: Users will be less likely to engage in selective exposure than the two baseline conditions when the attitude-inconsistent story is relevant and/or salient and the attitude-consistent story is not.

Data and methods

In two pre-registered (https://osf.io/mxnyr and https://osf.io/wjz8x) between-subject experiments with Norwegian Internet users, I manipulated an online choice environment. The respondents in both experiments were exposed to manipulated news content on a news site that they could browse freely. The news site was embedded in a larger online time-sharing survey panel, the research-driven Norwegian Citizen Panel (NCP), and featured a front page with news headlines that participants could click on to read the full story and two clearly labeled buttons for returning to the front page. The sample consisted of probability samples in which the entire Norwegian adult population had a known and equal chance of participation.

As a case for an online news environment, I simulated the mobile version of the interface of the authentic news site NRK.no—the online news site of the Norwegian public broadcaster the NRK (which, according to Alexa.com, is the sixth most popular website and the second most popular news site in Norway in 2021). The respondents had 2 minutes to browse the news site before the survey continued automatically. This time limit was set to encourage the respondents to be selective, and the NCP's pilot interviews indicated that respondents did not have sufficient time to read all articles. According to Alexa.com (per 10.6.21), NRK.no has an average daily time on site of around 2–2.5 minutes, which indicates that a browsing time of 2 minutes is realistic. Because the NRK is a public broadcaster, none of the news articles on the NRK's site require a subscription to be accessed, meaning that all adult Norwegians should have an equal opportunity to access the website and its news content. Note also that while a range of Norwegian news sites are already using NRSs to personalize the user experience on their front pages, the front page of the NRK.no is not (yet) among them. The NRK participated in piloting the study and provided explicit consent to use their logo in this experiment.

Data

I collected the data for this experiment in February 2021 (attitude items for Experiment 1), June 2021 (Experiment 1), and

December 2021 (attitude items and Experiment 2) through the 20th, 21st, and 22nd panel waves of the NCP. A total of 3,263 respondents who had responded to an attitude question on immigration in February 2021 (the 20th panel wave of the NCP) participated in Experiment 1. About 3,105 first-time recruited respondents participated in Experiment 2. None of the participants in Experiment 2 participated in Experiment 1. I conducted two experiments in separate waves to learn from the first experiment and, as a result of learning from the first experiment, tested the hypotheses toward a more ecologically valid baseline condition in the second experiment. The respondents for both experiments were gathered through the postal recruitment of 20,000 individuals over 18 years of age, randomly selected for recruitment from Norway's national registry: a list of all individuals who either were or had been residents in Norway.

Stimulus material and treatments

Across all conditions in both Experiments 1 and 2, the interface displayed eight authentic news stories published on NRK's news site or similar Norwegian news sites during the winter of 2021. All stories were shortened and anonymized (e.g., removed names mentioned in the story) to function as stimuli. Each story was between 92 and 143 words long ($M = 128.1$, $SD = 8.23$). Some articles were longer than others in some conditions—see the description of treatment conditions below—due to adding information about a specific city or county in the title.

Across all conditions in both experiments, each story featured its own picture to illustrate the stories (the same picture was used both on the front page and in the full article). Following the NRK interface, the stories were ordered vertically, with only one article per row on the front page. Two of the stories focused on immigration (one positive story and one negative story). The remaining news stories were filler stories; that is, they were not political, did not address the issue of immigration or politics, and did not feature negative or positive information about immigration. The design choice of including two stories on the same issue in the interface ensured that the issue was not the driving force for selection. Although both stories featured information about immigration, they described different topics within the main topic of immigration. In Experiment 1, a random subsample ($N = 1{,}790$) was asked to evaluate the valence of either the negative or positive immigration story on a scale from 1 ("very negative") to 5 ("very positive") (*without* any information about a geographical place). A t test showed that the negative story was statistically significantly ($p < 0.001$) more likely to be evaluated as negative toward immigration ($M = 2.53$, $CI = [2.48, 2.58]$) and the positive story was more likely to be evaluated as positive toward immigration ($M = 3.58$, $CI = [3.53, 3.63]$). This difference and pattern were significant regardless of the respondents' immigration attitudes.

Treatment and conditions

Across all conditions in both Experiments 1 and 2, the front page showed two articles on immigration (one positive and one negative) and six filler articles, a total of eight articles.

Experiment 1

In Experiment 1, participants were randomly assigned to one of nine conditions—a 3 (*Relevance treatment conditions*: geographically relevant attitude-consistent article vs. geographically relevant attitude-inconsistent article vs. relevance not present) × 3 (*Salience treatment conditions*: salient attitude-consistent article vs. salient attitude-inconsistent article vs. salience not present) design. The baseline condition did not include a salience or a relevance treatment and showed *all* eight articles in random order with equal size on the front page of the news site and with no mention of a geographic place in the headline. In the salience treatment conditions, the order of the immigration articles was fixed: either the negative immigration story was ordered on the top and the positive immigration article was ordered at the bottom of the front page, or vice versa. The story that was ordered on the top also featured an increased picture size. In the relevance treatment conditions, I followed the approach by Nanz and Matthes (2020) for manipulating relevance and included information about the city (if the respondent lived in one of the four largest cities in Norway) or county in which the respondent lived. Information about the city or the county was included in all versions of the negative and positive immigration stories (including the baseline conditions) on the news story level and was randomly included in the headline of either the negative immigration story, the positive immigration story, or not included at all. This design decision simulates an NRS that understands that the main text in the full article mentions a place, that this place is geographically close to the user and thus should be relevant for the user, and that then highlights the place mentioned on the front page as a relevance recommendation for the individual user. Such location-dependent news recommendations based on, for instance, the GPS location of the user is already being tested in Norway, and the NRK produces news from several local offices spread around Norway, meaning that they could, in a counterfactual scenario, develop an NRS that identifies and highlights whether some information in the main text of a news story is geographically relevant to a user. Eighty percent of a random subsample ($N = 1{,}419$) in Experiment 1 indicated that the place mentioned was either "very close or they lived there" or "somewhat close," indicating that the relevance manipulation displayed a place in close geographical proximity to the respondents. As argued by Nanz and Matthes (2020), the benefit of manipulating relevance via geographical proximity is that it is hardly confounded with political variables.

Experiment 2

In Experiment 2, participants were randomly assigned to one of four conditions—two baseline conditions (human editor-determined order baseline vs. randomized order baseline) and two treatment conditions (salient and geographically relevant attitude-consistent article vs. salient and geographically relevant attitude-inconsistent article). The two baseline conditions in Experiment 2 presented (a) the eight stimulus articles in random order (i.e., identical to the baseline in Experiment 1) or (b) a close-to-authentic baseline in which the public service broadcaster NRK.no's online front-page editors determined the order of the eight stimulus articles. To construct the second baseline, Qualtrics survey software was used to survey the NRK's front-page editors ($N = 6$). An email with a link was sent from the leader of the front-page editors in the NRK to an email list of 21 people who worked on the front page in some way or another. The response rate was 29% (see the Supplementary Material for more details). The six participants ranked the same article headlines as the articles I used as stimuli, and the results of these rankings were used to determine the ranking (i.e., the order in which the articles were presented) in the second baseline. As the rankings by the six front-page editors varied, the respondents who were randomly assigned to the second baseline randomly received one of the six possible rankings and thus viewed the articles on the front page in the exact same order as one of the six front-page editors had determined.

Similar to Experiment 1, the treatment conditions in Experiment 2 either promoted the negative or positive immigration story through relevance and salience recommendations. The treatments in Experiment 2 were close to identical to the treatments in Experiment 1. However, contrary to Experiment 1, the two treatment conditions in Experiment 2 did not separate the treatment effects of salience from the treatment effects of relevance but instead focused on the *combined* effects of relevance and salience recommendations to increase statistical power by reducing the number of treatment groups from eight to two. See the Supplementary Material for more details on each treatment condition in Experiments 1 and 2.

Measures

The key dependent variable for both Experiments 1 and 2 is article click (0–1, $M = 0.31$, $SD = 0.32$). In the Supplementary Material, I also display the main result by reading time. To measure the effect of attitude consistency on article click (i.e., selective exposure), I coded the positive and negative headlines as either "attitude-consistent" or "attitude-inconsistent" by matching the story with the respondents' immigration attitudes. I used the question "Generally speaking, how advantageous or disadvantageous is it that immigrants come to live in Norway?" anchored on a seven-point scale from 1 ("a very great disadvantage") to 7 ("a very great advantage") ($M = 4.58$, $SD = 1.39$). In Experiment 1, the attitude question was fielded in the 20th panel wave. In Experiment 2, the attitude question was asked at the beginning of the survey. The matching between the articles and the attitude question was coded as a binary and continuous independent variable. The binary variable was coded as 0 ("attitude-inconsistent") and 1 ("attitude-consistent"), using 1–3 and 5–7 as the cut-off points. The continuous variable used the entire scale of the attitude question (1–7). In the main text, I focus on the binary measure, but the results are robust if I use the continuous variable instead (see the Supplementary Material).

Analysis

As indicated by the theoretical estimand introduced above, the unit-specific quantity is a news article. Accordingly, in both Experiments 1 and 2, I had data for clicking behavior and reading time for each of the eight articles separately. This allowed me to rearrange the data; thus, I had multiple observations per respondent. I display the results from both experiments separately and in a pooled analysis (controlling for the fixed effects of the experiments). To ensure consistency and comparability between the two experiments, I follow the strategy from the pre-registration of the second experiment: I use multilevel GLM (logistic, Bernoulli) models, as I have several observations of article click nested within each respondent. The analyses are displayed graphically but the full tables are available in the Supplementary Material.

Results

If the baseline theoretical hypotheses (H1a and b) are correct, we would see that respondents are more likely to click on the attitude-consistent article compared to the attitude-inconsistent article if we restricted the analysis to one of the baseline treatment conditions. Across all three baseline groups in the two experiments, I find that respondents are significantly more likely to select the consistent rather than the inconsistent article (Experiment 1: $b = 0.43$, $SE = 0.19$, $p < 0.05$; Experiment 2_{random}: $b = 0.56$, $SE = 0.14$, $p < 0.001$; Experiment 2_{editors}: $b = 0.65$, $SE = 0.14$, $p < 0.001$), supporting H1a and b. Figure 2 plots the marginal predicted mean of attitude consistency (the attitude-consistent article vs. the attitude-inconsistent article) on article click per experiment and condition, which graphically shows whether there was an increased likelihood of selecting the consistent article, a decreased likelihood of selecting the inconsistent article, or both. For all baseline groups, the likelihood of selecting the attitude-consistent article (Experiment 1: $M = 0.61$, $SE = 0.03$; Experiment 2_{random}: $M = 0.55$, $SE = 0.02$; Experiment 2_{editors}: $M = 0.59$, $SE = 0.02$) was statistically significantly higher than the grand mean (0.5). In Experiment 1, the selection of the attitude-inconsistent article was not significantly different from the grand mean ($M = 0.53$, $SE = 0.03$), and in Experiment 2, the attitude-inconsistent article was significantly less likely to be selected, compared to the grand mean, in the random condition ($M = 0.45$, $SE = 0.01$) but not in the editor condition ($M = 0.48$, $SE = 0.02$). Figure 2 also shows that in the condition that promotes the attitude-inconsistent article with both relevance and salience recommendations (Figure $2_{\text{Experiment1(g)}}$ and Figure $2_{\text{Experiment2(d)}}$), the respondents were significantly *more* likely, compared to the grand mean, to click on the attitude-inconsistent article ($M = 0.56$, $SE = 0.03$; $M = 0.55$, $SE = 0.02$, respectively).

The RISE framework not only expected selective exposure to occur but also that the chance should be increased (H2) or decreased (H3) compared to the baseline conditions, depending on whether an NRS promoted attitude-consistent article or the attitude-inconsistent article through relevance and/or salience recommendations. These hypotheses were examined by regressing attitude consistency and the treatment conditions on the article click, including an interaction term between attitude consistency on the one side and the treatment conditions on the other side of the interaction term. The results are illustrated in Figure 3, which plots the interaction effects using the random baselines as the reference category. In line with H2, the pooled result displayed in Figure 3 shows that respondents were significantly *more* likely than in the random baseline condition to click on the attitude-consistent article compared to the attitude-inconsistent article when the attitude-consistent article was promoted through a relevance recommendation and/or salience recommendation. In line with H3, the pooled result displayed in Figure 3 shows that the respondents were significantly *less* likely than the baselines to click on the attitude-consistent article compared to the attitude-inconsistent article when the attitude-inconsistent article was promoted through a relevance recommendation and/or salience recommendation. Experiment 1 adds important nuances. In the conditions that promote the attitude-consistent article in Experiment 1, I do not find substantively different effect sizes in terms of increasing selective exposure when including only relevance recommendations ($b = 0.6$,

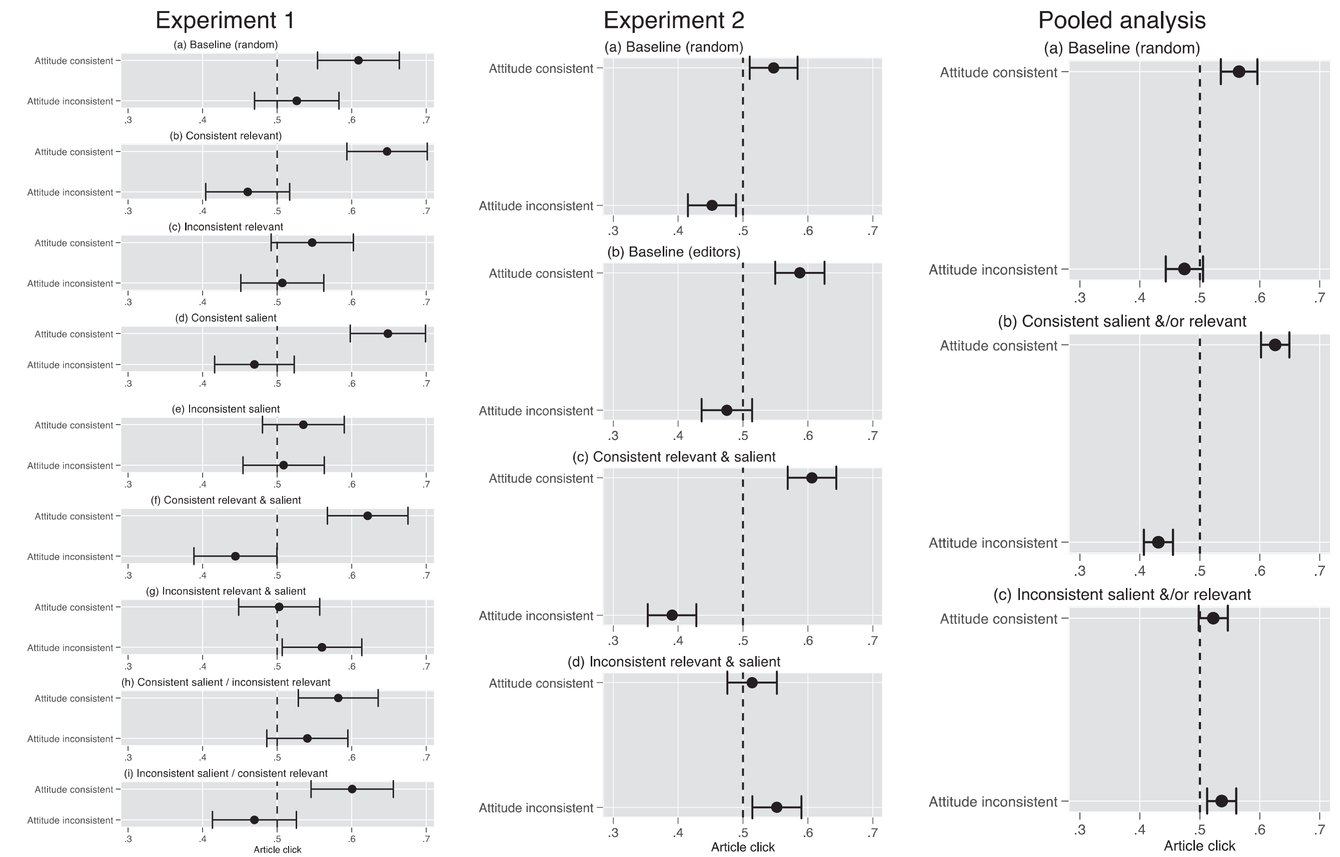

Figure 2. Marginal predicted means of the selection of attitude-consistent and attitude-inconsistent news articles by treatment condition.

Note. The dots represent the point estimates of the effects. The bars show 95% confidence intervals. Results of the GLM multilevel models. The dependent variable is article click (0–1). The figure shows the results from Experiment 1 ($N = 2{,}822$), Experiment 2 ($N = 2{,}638$), and in a pooled analysis of Experiments 1 and 2 ($N = 4{,}200$), controlling for the fixed effects of the experiments.

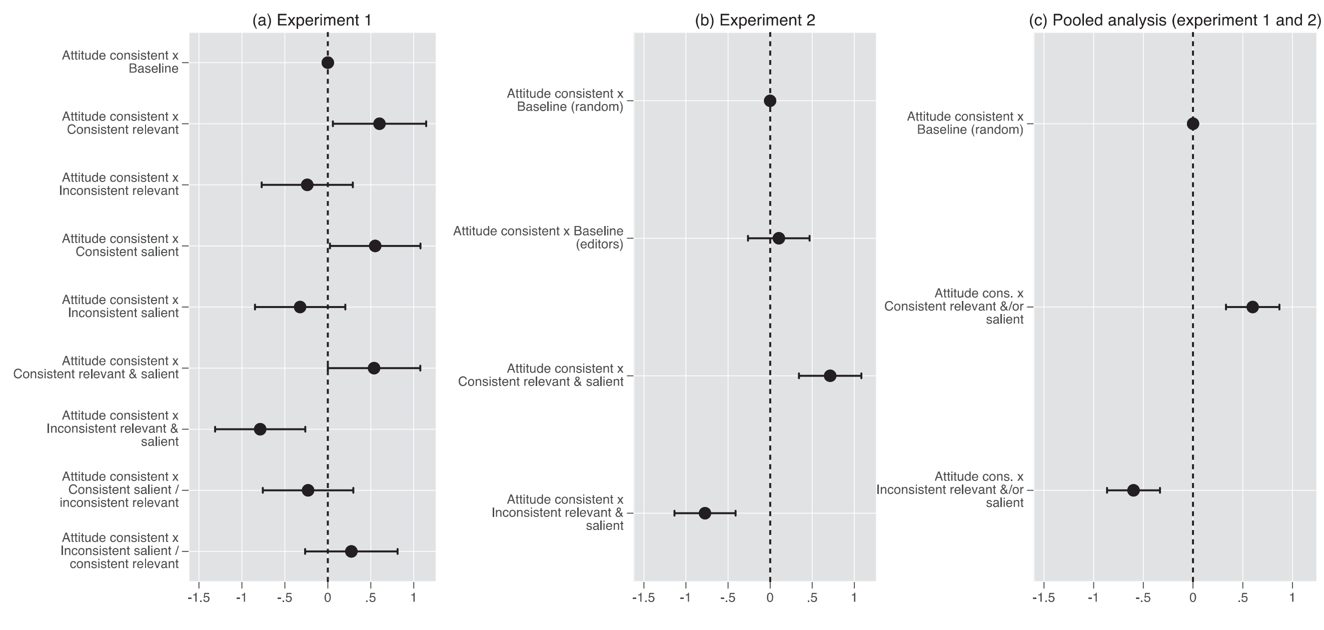

Figure 3. Interaction effects between attitude consistency of news articles and treatment conditions.

Note. The dots represent the point estimates of the effects. The bars show 95% confidence intervals. The dots without a bar represent the reference category. Results of the GLM multilevel models. The dependent variable is article click (0–1). The figure shows the results from Experiment 1 ($N = 2{,}822$), Experiment 2 ($N = 2{,}638$), and in a pooled analysis of Experiments 1 and 2 ($N = 4{,}200$), controlling for the fixed effects of the experiments.

$SE = 0.28$, $p < 0.05$), only salience recommendations ($b = 0.55$, $SE = 0.27$, $p < 0.05$), or combining both treatments ($b = 0.54$, $SE = 0.27$, ns). However, in the conditions that promote the attitude-inconsistent article in Experiment 1, I find substantively different effect sizes in terms of decreasing selective exposure. When only the relevance recommendation ($b = -0.24$, $SE = 0.27$, *ns*) or only the salience recommendation ($b = -0.32$, $SE = 0.27$, *ns*) is present, I do not find a significant difference. Yet, when both the relevance and the salience recommendation are present ($b = -0.79$, $SE = 0.27$, $p < 0.01$), I find a substantively larger and significant negative effect compared to the random baseline condition (i.e., decreased selective exposure). Figure 3 also shows that, while there were condition-dependent differences in the chance of selecting the attitude-inconsistent article, there were no significant differences between the two baseline treatments in Experiment 2 in terms of the difference between selecting the attitude-consistent and inconsistent article ($b = 0.1$, $SE = 0.19$, *ns*).

The number of respondents ($n \approx 650$) in each treatment group in Experiment 2 also allowed for analyses of heterogeneous effects. I do not find substantive differences in terms of immigration attitude, age, gender, interest in politics, immigration issue importance, or the use of smartphones to answer the survey (see the Supplementary Material for details).

Discussion

NRSs are increasingly and seamlessly used in online news environments. While prior research has devoted considerable attention to the assumption that selective exposure will increase through the causal variable of system-driven customization, the results are mixed (Arguedas et al., 2022; Terren & Borge-Bravo, 2021). The conditions under which NRSs influence or have no effect on selective exposure remain an unresolved puzzle. This article has investigated a critical aspect of this puzzle by introducing the RISE framework for modeling how news exposure decisions are influenced when NRSs are either programmed to increase or decrease the chance that selective exposure occurs.

The RISE framework and the empirical analysis aim to help clarify when selective exposure behavior is not only present in online news environments but is also causally increased or decreased by an NRS. This brings the field one step closer to solving the puzzle of the conditional effects of NRSs on selective exposure. As argued by Helberger (2019), because NRSs are programmed by human beings, their effects should be conditional on human design decisions and their values. The RISE framework aims to outline how such design decisions can influence selective exposure. I find that if simulated NRSs systematically highlight some content and downplay other content through relevance and salience recommendations, users are likely to follow such recommendations. If NRSs highlight attitude-consistent content, users are *more* likely to engage in selective exposure. If NRSs highlight attitude-inconsistent content, users are *less* likely to engage in selective exposure. These effects do not seem to be masked by the possible effects of the *absence* of human editorial decisions, as I find that simply randomizing the order of news stories through an automated system produces overall effects on selective exposure that are on par with a choice environment shaped by human editors. Future developers of NRSs should, however, be aware of an important differentiation of such selective exposure mechanisms; in line with Garrett (2009), I find that selective exposure to attitude-consistent articles is not necessarily equivalent to selective avoidance of attitude-inconsistent articles.

These findings have implications for an ongoing scholarly debate about algorithmic customization technology's impact on democracy—particularly the deliberative democracy model (Helberger, 2019). While some argue that NRSs increase selective exposure (Dylko, 2016; Pariser, 2011), others argue that NRSs have no notable causal effect on selective exposure (Bruns, 2019b; Dahlgren, 2021; Zuiderveen Borgesius et al., 2016) or can decrease selective exposure (Helberger et al., 2018). This study illustrates that all three arguments may be true, depending on what the NRS is designed to achieve. As such, it empirically illustrates Helberger's (2019) argument that the democratic implications of NRSs are partly determined by design decisions in terms of what different NRSs are programmed to achieve, which transfers the responsibility for the democratic implications of NRSs from the technology itself to the decisions surrounding the implementation and design of the technology. This study also illustrates the benefits of modeling counterfactual scenarios of the democratic implications of NRSs. When studying the NRSs that are currently used by social media platforms or news sites, researchers are often dependent on prime access to such technology—access that social media platforms or news sites do not necessarily share with researchers. The beauty of counterfactuals is that scholars can bypass these restrictions and gain causal evidence on the conditions under which NRSs are, for the better or the worse, for democracy. I encourage future research on the democratic implications of NRSs to make use of the RISE framework and further explore the input and output sides of the framework. For instance, the RISE framework is relevant for studies that aim to empirically test the conceptual framework of Helberger (2019) to assess the democratic threats and opportunities of NRSs. Helberger's (2019) framework highlights that the democratic consequences of NRSs depend on the democratic model one applies to assess such consequences.

This article also makes a noteworthy methodological contribution in terms of increasing the ecological and external validity of selective exposure research. While prior studies have unobtrusively captured selective exposure on an experimentally controlled news site and used nationally representative probability samples, this study also includes authentic human editors to determine the order of the news stories to establish an authentic baseline and goes beyond the often criticized single-experiment studies by including two closely related experiments. The design choice of including six nonpolitical filler stories also comes closer to real news selection environments than most prior studies on selective exposure, which typically feature only a choice between political news (Feldman et al., 2013).

While this article primarily discusses NRSs' influence on selective exposure to political information, the findings also have significant implications for the broader field of communication research. NRSs not only promote political news content but also content such as advertising, entertainment, and health information (Lazer, 2015). The RISE framework is not exclusively relevant to political content. It could also be used to study the conditions under which people share algorithmically recommended (mis)information online or decide to follow health advice from the government. Although I have focused on online news sites, the RISE framework should also be of direct relevance for research on NRSs on social media platforms in which NRSs might also influence and be influenced by one's contacts and network, as well as for ongoing policy debates on algorithmic accountability, such as regulating and increasing the transparency of algorithms (Diakopoulos, 2019).

Five key limitations of this study warrant further discussion. First is the design choice of including two stories on immigration to hold the topic constant comes, at least to some extent, at the expense of ecological validity, as it is less likely (yet far from unlikely) that an online news site simultaneously presents stories from different angles and within different subtopics on the same topic on the front page. Relatedly, as this initial test of the RISE framework focused only on the immigration issue for measuring selective exposure, it cannot necessarily generalize to other topics that spark political disagreement. Future applications of the framework should consider varying several topics. Thirdly, although relevance and salience recommendations are highly plausible factors that one could use to influence selective exposure through an NRS, a range of other contextual factors could possibly yield stronger effects. Future studies could seek to systematically test whether some factors yield more pronounced effects than others. Fourthly, while I have contributed evidence of the immediate effects of simulated NRSs programmed to influence selective exposure, the extent to which NRSs can reinforce feedback loops over time requires further empirical attention. A key assumption in the literature arguing for causal effects of NRSs on selective exposure is that selective exposure behavior will increase over time until, as Slater (2007) puts it, "a satisfactory equilibrium is reached" (p. 373) through a negative feedback loop, as "selectivity of attitude- and identity-consistent content are likely to operate only to the extent necessary to maintain a reasonable level of comfort with respect to protecting identity-central attitudes and beliefs" (p. 375). Conversely, the selection of attitude-inconsistent content should increase and subsequently be reinforced until an equilibrium is reached, if NRSs are programmed to decrease selective exposure. I encourage future work to empirically test and expand the RISE framework to include longitudinal conditional causal effects of NRSs on selective exposure. Finally, it is important to emphasize that the empirical approach in this study was an important first step, using a "classic" method of testing the effects on selective exposure. The NRS tested in this study were not authentic; rather, the results of experimentally varied choice environments displayed news articles as if they had been shaped by the NRS. I encourage future applications of the RISE framework to build authentic NRSs to test the input and output sides of the RISE framework. While the industry, to the best of my knowledge, has yet to develop NRS prototypes that allow for a rigorous empirical investigation of the conditions under which NRSs increase or decrease the chance that selective exposure occurs, the academic community has created software that allows for studies of NRSs' influence on exposure diversity (but not effects on selective exposure) (e.g., Heitz et al., 2022; Joris et al., 2022; Loecherbach & Trilling, 2020). Future authentic applications of the RISE framework have two so-called "cold start problems" that they need to overcome. The first is to overcome the "new user problem" (Rubens et al., 2015), that is, to learn the user's preferences regarding a range of political issues. The second problem is the "new item problem" (Musto et al., 2022), that is, to automatically determine whether news content would be consistent or inconsistent with this preference without data about how others have interacted with the news content. While laborious and demanding in terms of resources and technological development, possible solutions to these problems could be (a) through active learning (Rubens et al., 2015), that is, that the users actively provide answers on their preferences through a quick survey, and (b) through improvements in content-based recommendations where NRSs rely on automated content analysis in order to categorize its topic and framing of an issue (e.g., whether an article is clearly pro- or anti-immigration).

A natural next step will be to study and review moderators of selective exposure that are relevant for NRSs to promote and to build prototypes of NRSs that implement such factors and that can be tested in authentic environments. I particularly encourage future research to study technical solutions for programming NRSs (see, e.g., Joris et al., 2022) to influence selective exposure and its effects, as well as to study the conditions under which moderators of selective exposure are likely to correlate with the different optimalization goals of NRSs, such as increasing audiences' engagement with content. Such insights would help bridge the gap between the counterfactual scenarios described in the RISE framework and the NRSs currently used by different news sites, platforms, and services.

Supplementary material

Supplementary material is available online at *Journal of Communication* online.

Data availability

The data underlying this article is available free of charge for scholars at the NSD—Norwegian center for research data: https://search.nsd.no/series/ed271b1c-2595-47e4-8c97-3fcc00f02368.

Funding

This work was supported by industry partners and the Research Council of Norway with funding to MediaFutures: Research Centre for Responsible Media Technology and Innovation, through The Centres for Research-based Innovation scheme, project number 309339. The author is grateful that the MediaFutures: Research Centre for Responsible Media Technology & Innovation funded this study and made the innovative close cooperation with the NRK possible. This work was also supported by the Research Council of Norway with funding to the NEWSREC project, project number 324835.

Conflicts of interest: None declared.

Acknowledgments

The author thanks the NRK for allowing him to use their logo in the experiment and for distributing the survey to the NRK front page editors, which was crucial for increasing the study's ecological validity. He is also indebted to Mikael Poul Johannesson who provided crucial help with preparing the data for analysis and for sharing ideas on how to analyze the data. He is also grateful for Øivind Skjervheim and his team at Ideas2Evidence for devoting considerable time and effort in designing impressive technical solutions for the experiment. This project would not have been possible without the help and critical minds of both Mikael and Øivind. He also thanks the Editorial team and the anonymous reviewers for *Journal of Communication*, the anonymous reviewers for the ICA, Hallvard Moe, Morten Skovsgaard, Damian Trilling, Mehdi Elahi, participants at the NoPSA22 Political Communication workshop, as well as participants at the University of Bergen's Media Use Research group, for helpful comments, suggestions, and discussions.

References

Arguedas, A. R., Robertson, C., Fletcher, R., & Nielsen, R. (2022). Echo chambers, filter bubbles, and polarisation: A literature review. *Reuters Institute for the Study of Journalism*.

Atkin, C. (1973). Instrumental utilities and information seeking. In P. Clarke (Ed.), *New models for mass communication* (pp. 205–242). SAGE.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Beam, M. A. (2014). Automating the news: How personalized news recommender system design choices impact news reception. *Communication Research*, *41*(8), 1019–1041. https://doi.org/10.1177/0093650213497979

Bechmann, A., & Nielbo, K. L. (2018). Are we exposed to the same "news" in the news feed? An empirical analysis of filter bubbles as information similarity for Danish Facebook users. *Digital Journalism*, *6*(8), 990–1002. https://doi.org/10.1080/21670811.2018.1510741

Bozdag, E., & van den Hoven, J. (2015). Breaking the filter bubble: Democracy and design. *Ethics and Information Technology*, *17*(4), 249–265. https://doi.org/10.1007/s10676-015-9380-y

Bruns, A. (2019a). After the 'APIcalypse': Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, *22*(11), 1544–1566. https://doi.org/10.1080/1369118X.2019.1637447

Bruns, A. (2019b). *Are filter bubbles real?* John Wiley & Sons.

Dahlgren, P. M. (2021). A critical review of filter bubbles and a comparison with selective exposure. *Nordicom Review*, *42*(1), 15–33. https://doi.org/10.2478/nor-2021-0002

Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Harvard University Press.

Dylko, I. (2016). How technology encourages political selective exposure. *Communication Theory*, *26*(4), 389–409. https://doi.org/10.1111/comt.12089

Dylko, I., Dolgov, I., Hoffman, W., Eckhart, N., Molina, M., & Aaziz, O. (2017). The dark side of technology: An experimental investigation of the influence of customizability technology on online political selective exposure. *Computers in Human Behavior*, *73*, 181–190. https://doi.org/10.1016/j.chb.2017.03.031

Dylko, I., Dolgov, I., Hoffman, W., Eckhart, N., Molina, M., & Aaziz, O. (2018). Impact of customizability technology on political polarization. *Journal of Information Technology & Politics*, *15*(1), 19–33. https://doi.org/10.1080/19331681.2017.1354243

Feldman, L., Stroud, N. J., Bimber, B., & Wojcieszak, M. (2013). Assessing selective exposure in experiments: The implications of different methodological choices. *Communication Methods and Measures*, *7*(3–4), 172–194. https://doi.org/10.1080/19312458.2013.813923

Fletcher, R., Kalogeropoulos, A., & Nielsen, R. K. (2021). More diverse, more politically varied: How social media, search engines and aggregators shape news repertoires in the United Kingdom. *New Media & Society*. https://doi.org/10.1177/14614448211027393

Fletcher, R., & Nielsen, R. K. (2017). Are news audiences increasingly fragmented? A crossnational comparative analysis of cross-platform news audience fragmentation and duplication. *Journal of Communication*, *67*(4), 476–498. https://doi.org/10.1111/jcom.12315

Garrett, R. K. (2009). Politically motivated reinforcement seeking: Reframing the selective exposure debate. *Journal of Communication*, *59*(4), 676–699. https://doi.org/10.1111/j.1460-2466.2009.01452.x

Garrett, R. K., & Resnick, P. (2011). Resisting political fragmentation on the Internet. *Daedalus*, *140*(4), 108–120. https://doi.org/10.1162/DAED_a_00118

Haim, M., Graefe, A., & Brosius, H.-B. (2018). Burst of the filter bubble? Effects of personalization on the diversity of Google news. *Digital Journalism*, *6*(3), 330–343. https://doi.org/10.1080/21670811.2017.1338145

Hart, W., Albarracın, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*(4), 555–588. https://doi.org/10.1037/a0015701

Hassell, H. J., Holbein, J. B., & Miles, M. R. (2020). There is no liberal media bias in which news stories political journalists choose to cover. *Science Advances*, *6*(14), 1–8. https://doi.org/10.1126/sciadv.aay9344

Helberger, N. (2019). On the democratic role of news recommenders. *Digital Journalism*, *7*(8), 993–1012. https://doi.org/10.1080/21670811.2019.1623700

Helberger, N., Karppinen, K., & D'Acunto, L. (2018). Exposure diversity as a design principle for recommender systems. *Information, Communication & Society*, *21*(2), 191–207. https://doi.org/10.1080/1369118X.2016.1271900

Heitz, L., Lischka, J., Birrer, A., Paudel, B., Tolmeijer, S., Laugwitz, L., & Bernstein, A. (2022). Benefits of diverse news recommendations for democracy: A user study. *Digital Journalism*, 1–21. https://doi.org/10.1080/21670811.2021.2021804

Huszár, F., Ktena, S. I., O'Brien, C., Belli, L., Schlaikjer, A., & Hardt, M. (2022). Algorithmic amplification of politics on Twitter. *Proceedings of the National Academy of Sciences*, *119*(1), https://doi.org/10.1073/pnas.2025334119

Joris, G., Vercoutere, S., De Clercq, O., Van Damme, K., Mechant, P., & De Marez, L. (2022). Nudging towards exposure diversity: Examining the effects of news recommender design on audiences' news exposure behaviours and perceptions. *Digital Journalism*, 1–22. https://doi.org/10.1080/21670811.2022.2106445

Karimi, M., Jannach, D., & Jugovac, M. (2018). News recommender systems–Survey and roads ahead. *Information Processing & Management*, *54*(6), 1203–1227. https://doi.org/10.1016/j.ipm.2018.04.008

Knobloch, S., Carpentier, F. D., & Zillmann, D. (2003). Effects of salience dimensions of informational utility on selective exposure to online news. *Journalism & Mass Communication Quarterly*, *80*(1), 91–108. https://doi.org/10.1177/107769900308000107

Knobloch-Westerwick, S. (2015). *Choice and preference in media use: Advances in selective exposure theory and research*. Routledge.

Knobloch-Westerwick, S., Carpentier, F. D., Blumhoff, A., & Nickel, N. (2005). Selective exposure effects for positive and negative news: Testing the robustness of the informational utility model. *Journalism & Mass Communication Quarterly*, *82*(1), 181–195. https://doi.org/10.1177/107769900508200112

Lazer, D. (2015). The rise of the social algorithm. *Science*, *348*(6239), 1090–1091. https://doi.org/10.1126/science.aab1422

Levy, R. (2021). Social media, news consumption, and polarization: Evidence from a field experiment. *American Economic Review*, *111*(3), 831–870. http://doi.org/10.1257/aer.20191777

Lodge, M., & Taber, C. S. (2013). *The rationalizing voter*. Cambridge University Press.

Loecherbach, F., & Trilling, D. (2020). 3bij3: Developing a framework for researching recommender systems and their effects.

Computational Communication Research, 2(1), 53–79. http://doi.org/10.5117/CCR2020.1.003.LOEC
Loecherbach, F., Welbers, K., Moeller, J., Trilling, D., & Van Atteveldt, W. (2021). Is this a click towards diversity? Explaining when and why news users make diverse choices. In *13th ACM web science conference 2021*, *Virtual, 21–25 June*, (pp. 282–290). New York: Association for Computing Machinery. https://doi.org/10.1145/3447535.3462506
Lundberg, I., Johnson, R., & Stewart, B. M. (2021). What is your estimand? Defining the target quantity connects statistical evidence to theory. *American Sociological Review*, *86*(3), 532–565. https://doi.org/10.1177/00031224211004187
Mattis, N., Masur, P., Möller, J., & van Atteveldt, W. (2022). Nudging towards news diversity: A theoretical framework for facilitating diverse news consumption through recommender design. *New Media & Society*. https://doi.org/10.1177/14614448221104413
Mitova, E., Blassnig, S., Strikovic, E., Urman, A., Hannak, A., de Vreese, C. H., & Esser, F. (2022). News recommender systems: A programmatic research review. *Annals of the International Communication Association*, 1–30. https://doi.org/10.1080/23808985.2022.2142149
Musto, C., Gemmis, M. D., Lops, P., Narducci, F., & Semeraro, G. (2022). Semantics and content-based recommendations. In *Recommender systems handbook* (pp. 251–298). Springer.
Möller, J., Trilling, D., Helberger, N., & van Es, B. (2018). Do not blame it on the algorithm: An empirical assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society*, *21*(7), 959–977. https://doi.org/10.1080/1369118X.2018.1444076
Mummolo, J. (2016). News from the other side: How topic relevance limits the prevalence of partisan selective exposure. *The Journal of Politics*, *78*(3), 763–773. https://doi.org/10.1086/685584
Nanz, A., & Matthes, J. (2020). Learning from incidental exposure to political information in online environments. *Journal of Communication*, *70*(6), 769–793. https://doi.org/10.1093/joc/jqaa031
Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin UK.
Pearl, J. (2009). *Causality*. Cambridge University Press.
Rubens, N., Elahi, M., Sugiyama, M., & Kaplan, D. (2015). Active learning in recommender systems. In Ricci, F., Rokach, L., Shapira, B. (eds) *Recommender systems handbook* (pp. 809–846). Springer. https://doi.org/10.1007/978-1-4899-7637-6_24
Slater, M. D. (2007). Reinforcing spirals: The mutual influence of media selectivity and media effects and their impact on individual behavior and social identity. *Communication Theory*, *17*(3), 281–303. https://doi.org/10.1111/j.1468-2885.2007.00296.x
Terren, L., & Borge-Bravo, R. (2021). Echo chambers on social media: A systematic review of the literature. *Review of Communication Research*, *9 (2021)*, 99–118. https://doi.org/10.12840/ISSN.2255-4165.028
Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press. [Database].
Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age*. MIT Press.
Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review*, *5*(1), 1–16. https://doi.org/10.14763/2016.1.401