Out of control?

Using interactive testing to understand user agency in news recommendation systems

Moeller, J. ; Loecherbach, F.; Möller, J.; Helberger, N.

DOI
10.4324/9781003257998-11

Publication date
2023

Document Version
Final published version

Published in
News Quality in the Digital Age

8

OUT OF CONTROL?

Using interactive testing to understand user agency in news recommendation systems

Judith Moeller, Felicia Loecherbach, Johanna Möller and Natali Helberger

Introduction

Users of recommender systems on social media and beyond are thought to be involuntarily trapped to hear exclusively a perpetual echo of their existing thoughts and beliefs. In the words of Eli Pariser (2011), this constitutes a "personal ecosystem of information that's been catered by […] algorithms to who they think you are", the so-called "filter bubble." In the debate about this phenomenon, educating users and giving them more agency is often hailed as one of the key solutions to mitigate negative effects. Many policy makers hold the belief that explaining how algorithms function and giving people options to curate their own information will empower users and help them break out of these bubbles. The Digital Services Act which is the key regulatory framework of online information in the EU is a good example of this approach (Morais Carvalho et al., 2021). The underlying assumption is that users who understand the system and are presented with control options can exercise meaningful control to ensure the curated information stream fits their needs and reflects their interest.

Yet, this assumption has not yet been empirically tested. On the contrary, according to Guzman (2019), scholarship into automatic news dissemination generally lacks an understanding of the technology's communicative role from the point of view of the audience. The notion of empowered users is based on the potential control and theoretical agency that users hold in the context of recommender systems. Indeed, users who are willing and able can often exercise meaningful control over the news environment either through changing the settings of the recommender systems (explicit control), or by steering the system through their selection behavior in feedback loops (implicit control) (Bozdag, 2013), thus making use of the personalization algorithm employed by platforms to meet user

DOI: 10.4324/9781003257998-11

preferences by customizing newsfeeds based on past engagement with content. However, explorative studies in the field (Monzer et al., 2020) show that while users do perceive themselves as active participants when interacting with recommender systems, they also complain about a perceived lack of effective control options (perceived control) This chapter aims to provide insight into the interplay of algorithmic settings and perceived agency. In other words, we want to understand whether different levels of actual control translate into perceived control. Using a novel experimental design (Loecherbach & Trilling, 2020) that allows users to engage with a news recommendation system, we aim to gain insight into how different forms of news personalization are connected to perceived user control and satisfaction with algorithmic news selection.

Studying recommender systems in different fields

The role of the user as an active controller of news recommendation systems is currently insufficiently studied in the different fields that investigate news recommender systems: computer science and social sciences, including normative approaches. This is a direct consequence of the different epistemologies used to understand news consumption in the different fields. In computational and information sciences, which have been the dominant fields for studying (news) recommendation systems, such systems were initially treated as a mathematical issue. In this line of research, the user is conceptualized as the target of recommendations, characterized by stable preferences. This means users are "more the objects than the subjects of this code" (Dahlgren, 2018, p. 23). Recent work has expanded this approach to include measures of diversity, novelty, and serendipity (Karimi et al., 2018), based on a more dynamic image of the user. Another recent development is to move beyond the stable conceptualization of users and test recommenders in live environments (for example, Karimi et al., 2018, or Jugovac & Jannach, 2017). While these developments certainly contribute to a more nuanced understanding of how users interact with algorithmically selected news, there are still both methodological and theoretical shortcomings that speak to an oversimplified conceptualization of the user. The findings are often based on small sample sizes (10–20 respondents) not acknowledging the heterogeneity among users. Additionally, no tests of theoretical assumptions or frameworks are carried out, and the main focus remains on having users give live feedback to improve the algorithms' accuracy and precision (Goossen et al., 2011) as well as marketing-related measures such as click-through rates.

In other fields, the study of recommender systems as a research phenomenon is more recent. In light of the debates about filter bubbles (Pariser, 2011) and the rise of social networking sites as gateways to news, the phenomenon has attracted broader scholarly attention, especially in the social sciences. While these studies employ a more holistic view of the user as an actor situated in a social context, they rarely account for the specific and limited possibilities users have to exercise control over recommendation systems and the dynamics of implicit control through feedback loops. Scholars in the social sciences are mainly focused on testing mechanisms of selective exposure by using one-time selection procedures with content specifically designed for the experimental purpose (Go et al., 2014) and focusing on pro- vs. counter-attitudinal news content (Beam & Kosicki, 2014). Often news stories are presented in a questionnaire format rather than a (real or constructed) news website and remain in one-time selection settings that are not compared to a recommendation algorithm (Zhu & Lee, 2019).

A different strand of research focuses on the user experience and understanding of algorithmic selection. In a qualitative study, Bucher (2017) charted how users perceive the Facebook algorithm through tweets and interviews of 25 ordinary users. She finds that the imaginary of the algorithm shapes how users interact with the system. To some users, this includes the perception that the recommender system has a faulty image of them, an observation that is line with the findings of Monzer and colleagues (2020). Yet many other users are not even aware that news feeds are algorithmically curated (Powers, 2017) and according to a study by Min (2019), only about 13 percent of users are actively challenging algorithmic selection, beyond following or unfollowing certain accounts, whereas about 38 percent do not engage in any form of explicit or implicit control.

From a normative perspective, discussions about recommender systems in the news domain in large part rely on democratic ideals and normative assumptions regarding who should be the gatekeeper and selector of news (Zuiderveen Borgesius et al., 2016; Helberger, 2019). Users are often seen as passive users rather than active actors in the news making and curation process, which also had a lot to do with the one-to-many logic of traditional mass communication services. That being so, with the arrival of more interactive forms of communication and also the growing importance of data, the traditional mass media logic of seeing users essentially as eyeballs has been juxtaposed by the more active control paradigm in data protection law, creating the expectation of users as active controllers of their personal data, if only properly informed and offered the right choices (Hildebrandt, 2019). However, to date there is little empirical research into how to design for meaningful agency and if doing so will indeed result in an increased perception of agency and control from the perspective of users (Strycharz et al., 2020). What both communication science and information law add to the field of computational science is thus the "noisy" element of the recommendation process: the user.

The complicated function of user agency in news consumption

Yet, neither the conceptualization of the stable user as object of the recommender engine, nor that of empowered users, willing and able to shape the news recommender system once they possess enough information and literacy, capture the reality of what it means to be a news consumer in the datafied democratic society. The user is both more and less than either imaginary. To arrive at an integrated conceptualization of users, we thus need to take a step back to understand the factors that differentiate perspectives on how users exercise control when consuming news and what hinders them from doing so. First, we need to explore individuals' relations with media as technologies and their mechanisms of information distribution.

It is important to note that trusting journalistic gatekeepers with the selection of information is necessary and normal in democratic societies (Dahlgren, 2018). This trust is the natural twin sibling of knowledge, as opportunities for taking information and insights for granted enables citizens to cope with demanding information processing work. As this delegation of informational duties has been ritualized for decades, it can be argued that the confidence placed in journalism as gatekeepers is extended to the algorithmic gatekeepers in news recommender systems (see Monzer et al., 2020). That being said, journalistic gatekeeping, which historically took over responsibilities for the choice, verification, and contextualization of information as well as offering opportunities for engagement with news, itself now loses influence, as the internet offers myriad additional channels. Expressed in more positive terms, the news media landscape becomes increasingly "dynamic and varied through the mingling of different news actors" (Russell & Waisbord, 2017, p. 71). At the same time, audience research shows that media consumption is largely ritualized (Lee & Delli Carpini, 2010). Media users compose their media repertoires and contextualization strategies based on habits in their social surrounding (Thorson et al., 2018). News consumers "learn" to trust (or not) media as a systematic resource. Related habits, we assume, persist.

There is also a self-managing element in news consumption (as well as in other dimensions of informed citizenship) that calls for further attention (Livingstone, 2000). Skills, practices, and general trends of self-determined political communication as related to building political knowledge and collecting information were formerly organized within more or less institutionalized publics. Now they increasingly depend on individuals' abilities to understand and cope with technological environments (Hintz et al., 2019). Practices of delegating trust, thus, are called into question and individuals' capacities to understand, reflect on, and critically use technology (Kannengießer, 2020) are at stake. Yet these abilities to critically reflect on technology are not as developed for recommender systems. On the contrary, users' knowledge on algorithms is limited (Cotter & Reisdorf, 2020), and the underlying technology is perceived as too complex to control at least by some of the users (Monzer et al., 2020).

We thus can distinguish three reasons that complicate the motivation and ability to embrace the agency news recommender systems provide to users. First, news users are socialized to delegate the task of selecting relevant news to trusted institutions. This is fostered by the fact that, second, in many regards news consumption is a routine behavior, in which users follow a behavioral pattern, without the motivation to take on an active role to steer the process. Lastly, the complexity of algorithmic recommendation systems and the perceived unresponsiveness of the feedback loop or feeling mis-profiled by the system can lead to disengagement among users (Bucher, 2017; Min, 2019).

User involvement in recommender systems

So how can users exercise control? In general, we can distinguish two forms of control: implicit and explicit control. The notion of explicit and implicit control refers to different strategies of personalization, namely "customization" and "recommendation" (Beam & Kosicki, 2014)—similar to concepts such as explicit and implicit personalization (Bozdag, 2013). Explicit control leaves the control to the user (who has to provide explicit preferences, for example, by stating which topics they like, etc.), while implicit control only allows limited control through the feedback loop. Implicit and explicit can be seen on a continuum from having full control over what is recommended vs. being exposed to recommendations with minimal options to provide feedback. In practice, they are often combined in hybrid forms.

Recommender systems involving the user are often termed "interactive recommender systems". He et al. (2016) mention the main reasons for involving the user in the process: strengthening user involvement and compensating "for deficiencies in recommendation algorithms" (p. 12)—mostly seeing user control as means to increase the accuracy of the predictions yielded by feedback mechanisms. Indeed, insights from other research fields such as gaming studies, marketing, or educational studies support the view that enhanced user control can not only increase the accuracy of the system but also heighten involvement (McNee et al., 2003) and learning perceptions (Ku et al., 2016). However, there are also reasons beyond the optimization of the recommendation process that are of importance to consider here. For example, Sundar (2008), in discussing customizability in computer-mediated communication, states that "greater interactivity allows for greater assertion of one's presence" (p. 62). Thus, in the end, more user control gives back agency—and in the same realm increases transparency of how the recommender works. Additionally, not giving the user the agency to interfere with the recommender in case of erroneous recommendations or a mis-profiling of the user has an impact on the user beyond satisfaction with the system: It can be seen as a threat to first-person authority and self-knowledge (Gertler, 2017). Therefore, adding control mechanisms and opportunities to interfere to the interface can be seen as means to give back the agency to the person who knows best.

H1a: Explicit personalization (customization) leads to higher perceived user control than implicit personalization.

This does not, however, exclude that some parts of the recommender system should still allow for discovery and unexpected results—since in some cases "people might be interested in things that they did not know they were interested in" (Bozdag, 2013, p. 217). In case no clear preferences are formed yet (such as when people have little experience with or interest in news selection), prompting discoveries can add value to using a recommender system.

Generally, there are different ways in which algorithms can select the news recommended to the user. Usually, a profile of a user is made based on the content they selected in the past and interests they explicitly indicated. In collaborative recommender systems, this profile is matched with a similar user to find items of interest. In content-based systems, articles are selected based on how similar they are with the content the user had read in the past. In this study, we are focusing on content-based systems—and here an important choice is the feature selection, i.e., what information about the article is used for matching it with the user profile. One approach is to use meta-data like the author, source, or genre of a news article or source to find matches that are in line with previous choices of the user. One example of such a meta-data-based recommendation algorithm is a topic-based recommender that deducts a topic preference from user behavior and includes more articles of the same topic in future recommendation sets.

A different approach is automatically analyzing the vocabulary of news articles a user has engaged with in the past and recommending more articles that share linguistic features with these articles (called a similarity-based recommender). The advantage of this approach is that it encompasses many different features of language, for example, if a user has a preference for complex language or specific actors. It is thus more fine-tuned compared to a topic recommender. The disadvantage is that the algorithm is so complex that it is difficult for users to understand how their past choices and behavior have influenced the recommendation sets they are receiving. When being given explicit visual cues on how the information environment has changed—for example, bright color topic tags that change depending on previous selections—the reaction of the recommender to choices of the individual becomes more apparent. Especially since cognitive capabilities are limited, attention can always only be targeted selectively at parts of the information environment (Posner, 1994), often looking for easy visual cues to be detected (Bodenhausen & Hugenberg, 2009). If those cues show a reaction to past choices, it should be easier for users to detect options for indirect control within a recommender system. We thus expect that the higher simplicity of the topic-based recommenders provides users with a larger sense of control.

> H1b: Having a topic-based recommender leads to higher perceived user control than a similarity-based recommender

However, while it can be expected that enhanced user control has beneficial outcomes for both economic and normative aspects of news recommender systems, research in this domain remains rather scarce: in a general overview of studies involving interactive recommender systems, (He et al., 2016), 15 papers are mentioned that specifically designed for or tested controllability. However, none of the studies were about the news domain. Apart from this, as mentioned above, the user studies employed relied on very small sample sizes and remain exploratory in nature. Additionally, most of them have not yet been tested over a longer period of time.

A more specific user profile and a repeated selection of items beyond one-time selection processes are only part of how recommender systems are used in real life. It might, for example, be true that in the short term, customization and enhanced control increase the satisfaction of the user—while in practice people do not use those options due to limited willingness to invest time in customization. It takes mental effort to make customization choices over and over again—since it requires "active exercise of personal choice" (Kang & Sundar, 2013, p. 2273) and thus depletes mental resources. Using one's self-agency can thus also be exhausting, especially when doing it over a longer period of time, possibly leading users not to take advantage of customization options. Indeed, Chung (2008) states: "online audiences are not using interactive features extensively contrary to anticipation by media scholars and the news industry" (p. 672). Nonetheless, based on past research (predominantly in other domains), it can be expected that direct customization requiring active user involvement gives a higher sense of control to the user compared to personalization that relies on behavioral factors without explicit interference by the user.

H2: Explicit personalization leads to higher satisfaction with a system than implicit personalization.

H3: The effect in H2 is mediated by perceived user control such that explicit personalization positively influences perceived user control which positively influences satisfaction with the system.

Methods

To test the above-mentioned research questions,[1] we used a platform for testing an interactive news recommender system for period of ten days (Loecherbach & Trilling, 2020). It is based on an open-source Python application and presents the user with a web interface showing a selection of nine different news articles in tiles. For each article, the title and a short teaser are shown; additionally, the topic is indicated with a colored tag. The news articles are retrieved from several RSS feeds of different Dutch news providers (similar to strategies employed by Nguyen and Nguyen (2016), or Phelan et al. (2010)). This ensures that the articles presented are recent as well as actual news—a crucial aspect since "in contrast to other domains like movie recommendation, the relevance of news items can change very rapidly" (Karimi et al., 2018, p. 1204). Thus, the website is intended to resemble a news website in a controlled design (i.e., without distracting advertisements and additional cues such as pictures and other tags). Only a set of nine items is used per recommendation round, since the amount and placement of stories plays a crucial role in recommender systems (Jugovac & Jannach, 2017), following an inverted U-shape where too few and too many choice options negatively impact the satisfaction with the system. Additionally, having the tile system allows for displaying all elements on one page without scrolling, limiting positioning effects.

Usage of the website

After the initial questionnaire, participants are redirected to the website where an account needs to be created. During the registration phase, users are assigned to four different experimental conditions:

Random. This group saves as a baseline and control group: Every time the user visits the website, nine random news stories are presented. There are no options to customize the interface and no recommendation algorithm is used, thus the participant simply gets a random selection of current RSS feeds.

Similarity-based behavior recommender. In this condition, stories are selected based on how similar they are to the articles a user has read in the past. It is thus a classic content-based recommender that aims at finding similar items to the user profile. For each article the user has read before, the three most similar new articles are used as potential candidates for new suggestions.[2] Out of this pool of articles, a random selection is made.

Topic-based recommender. The second recommender system is again content-based. However, it does not match the vocabulary of the read stories but rather their topics, in that it relies on specific tags or keywords for recommending the news stories and is less bound to specific content read but rather the broad overall interests of a person. In case a person has a specific interest in one topic, they will be shown more from this topic even if the articles do not match vocabulary-wise (i.e., from different types of sports); when having a broad interest in many different topics, the recommender will reflect that broadness. This recommender more clearly "reacts" to the selection of a person in that the topic tags indicate the personalization at one glance.

Customization. This condition by default shows only random stories—unless the user decides to customize the interface. Without interference, no personalization takes place. The option to customize essentially picks up the topic-based recommender: The user can decide for up to three topics to be recommended more often. This is similar to making a static user profile as it used to be an option in the old interface of Google News (Jugovac & Jannach, 2017, p. 12)—however, it can be changed by the user at any time. A different number of stories is shown for each selected topic, depending on whether one (six recommended stories per topic), two (three stories), or three (two stories) favorite domains were selected. Therefore, direct consequences of the customization should be noticeable for the user.

In all personalization conditions (two recommenders and customization), three of the nine stories remain randomly selected to always offer the possibility to select non-personalized stories and thus change the outcomes of the recommender system even if no customization options are given. Thus, two-thirds of the stories are personalized, without indicating to the user which are and which are not.

Other interface elements

The overall interface additionally includes elements of gamification to ensure interaction with the system and keep participants entertained enough to continue using the website over a longer period of time: With every interaction, a number of points are collected (for logging in, reading articles, rating articles) with daily limits to avoid the situation where all points can be collected in one go. This system allows the user to estimate how much interaction is still needed until finishing the study. Lastly, other feedback elements such as reporting articles that were displayed incorrectly (to avoid low ratings not due to content) and a contact form are given.

Final questionnaire

After having gathered eighty interaction points and having logged in on at least seven different days, participants are prompted to fill in a final questionnaire. It includes scales for measuring perceived control (Knijnenburg et al., 2011), satisfaction with the system (Chen et al., 2011), and behavioral intention ("If we were to launch an official version of this website, how likely is it that you would use it again?," 7-point scale—see Figure 8.1).

Explorative addition

Apart from testing the preregistered hypotheses, this study also included an additional exploration of the question of whether actual control is the determining factor when it comes to news personalization, or whether the feeling of being in control (i.e., having explicit visible control options) is enough. So the larger question is not which type of control creates more user satisfaction, but whether it is really only the illusion of control that provides perceived agency and thus satisfaction with the system. This question has not gotten much scholarly attention, but it is crucial to understand to make sense of the causal process at play. For this reason, participants in the two recommendation conditions (topic-based and

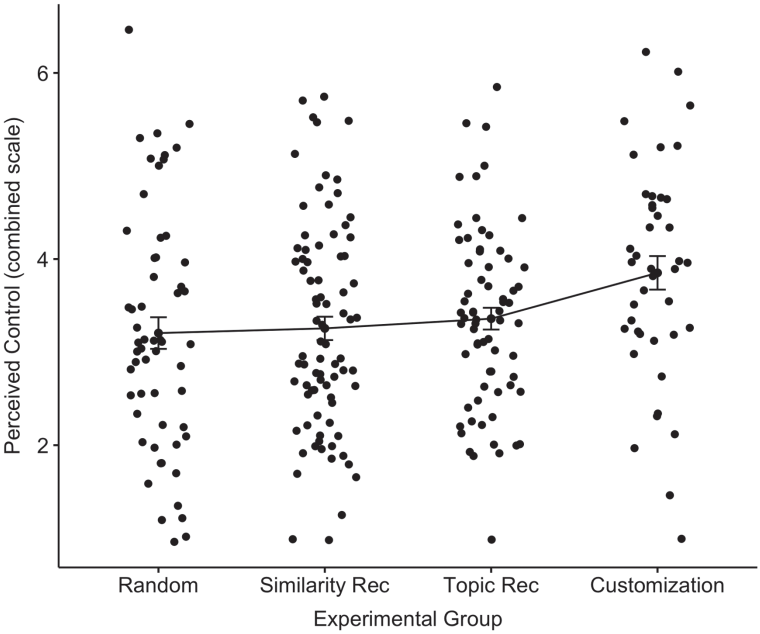

FIGURE 8.1 Perceived control (7-point-scale) in the different experimental groups.

similarity-based) were asked to continue the study for an additional three days and were being given additional control options they did not have before: All got a slider to control how many out of the nine stories on a page should be personalized. Additionally, users in the similarity-based recommender condition got a "diversity" slider to adjust how similar or dissimilar the recommendations should be to the content they read in the past. The users in the topic-based recommender condition got the same panel for indicating topic preferences that was given to the customization condition in the first part of the study. However, the control options actually worked only for half of the participants in this second part of the study—for the other half, they were just sliders without any real use, giving users the illusion that they were in control. After three days, participants were again asked for their feelings of control and satisfaction to be able to compare the impact of added controls to those variables and to see whether participants spotted if their control features did not work.

Results

The data collection took place between October and December 2019. In total, 1,753 users filled in the initial questionnaire, of those 1,585 agreed to proceed, 1,160 made an account on the website, and 1,029 activated their account by confirming their email. In total, 298 respondents qualified for the final questionnaire (enough interaction and days logged in). As specified in the preregistration, respondents who (1) had a standard deviation of 0 on one of the dependent variables, or (2) spent on average less than five seconds reading the news articles were excluded from the analysis, leading to a final sample of 248 respondents. Of those, 57 were in the random baseline group, 81 in the similarity recommender group, 68 got the topic recommender, and 42 could customize their settings. They are between 18 and 86 years old (M=46.19, SD=14.89), 54 percent identify as female and 53.6 %percent finished higher education. Randomization checks show no differences between any of the groups regarding socio-demographics or variables related to news consumption and political interest.

Of the 42 users who were given the option to customize their settings by selecting topics that they were particularly interested in, 16 did not use the option at all. Another twelve only used the option once and after that left the settings unchanged, meaning that two-thirds of respondents only sparsely made use of the control options given. The other users used the tool up to 16 times—however, again most of the activity happened within the first two days of signing up to the system. Additionally, users often inserted the same topic choices multiple times and did not change their preferences again.

Hypotheses testing

Looking at H1a and H1b, the differences between the groups regarding perceived control were examined. The seven items used to measure perceived control form a reliable scale (α=0.9). The items were thus averaged into one measure (M=3.37, SD=1.14). Since the normality assumptions are violated for perceived control, a Kruskal-Wallis test is used for analysis. A significant difference was found ($\chi 2(3, N= 248)=10.43$, $p=0.015$, $\eta 2=0.038$). Pairwise comparisons using the Wilcoxon rank sum test show significant differences between the customization group and all other groups (baseline p=0.017; similarity-based recommender, p=0.017; topic-based recommender, p=0.029). This shows that the customization condition with its explicit control indeed lead to higher perceived control among respondents compared to the baseline and the two recommender systems (implicit control). H1b received no support: the two recommender systems were not perceived differently in terms of controllability.

For H2, the two recommender groups are compared to the customization group regarding several satisfaction measures: explicit satisfaction (satisfaction with the website), implicit satisfaction (rating individual articles), and behavioral intention to return to the website. Since normality assumptions are violated, the Mann Whitney U test is used for comparison. The customization group compared to both recommendation groups combined does not show any higher explicit satisfaction (U=3501.5, p=0.24), similar results are found when comparing the customization to each recommender individually (topic recommender: U=1162.5, p=0.10; similarity-based recommender: U=1594, p=0.57). Regarding the ratings, again no differences could be found for comparing the customization group with the combined recommendation group (t(78.977)=0.21, p=0.83), the topic recommendation (t(95.845)=0.08, p=0.92) or the similarity-based recommendation (t(102.98)=-0.40, p=0.68). Lastly, regarding the behavioral intention for coming back to the website, no differences could be found for any of the comparisons (combined recommenders: U=3609.5, p=0.12; topic recommender: U=1191.5.5, p=0.14; similarity-based recommender: U=1457, p=0.18). Overall, this shows that no differences between the customization conditions and the recommender conditions could be found regarding any of the satisfaction measures, lending no support to H2. Since the direct effect could thus not be established, the mediation proposed in H3 was not tested.

Control vs. the feeling of being in control

To explore whether the illusion of control or actual control are related to satisfaction with the system, we analyzed how measures of satisfaction and perceived control changed after adding responsive or unresponsive control options. We used a mixed model ANOVA with perceived control and satisfaction as dependent variables and with the measurement point (t1, t2) as within factor and experimental group (topic recommender or content-based recommender), as well as control (fake or real) as between factors. For perceived control, no significant results for any of the variables or interactions could be found. The results for satisfaction show that while there are no main effects of any of the independent variables (group: F(1, 129)=0.196, p=0.659, control: F(1,129)=1.370, p=0.224, measurement: F(1,129)=0.081, p=0.776), a small significant interaction effect of measurement point and control can be found (F(1,129)=6,440, p=0.012). As Figure 8.2 shows, respondents who received actual control gained slightly in satisfaction between t1 and t2, while those who got non-working sliders showed a decline in satisfaction. In other words, when control is promised but not realized, users become more dissatisfied with the recommender system.

Discussion

In this chapter, we set out to understand how and if the possibility to control a news recommender system translates into perceived control and agency. From the perspective of engineers developing recommender systems, the user is conceptualized (1) as an entity that has to be modeled, that has to be represented in a mathematical way to select fitting items to it, or (2) as a consumer that needs to be satisfied and come back to the website frequently. In this sense, users are seen as bundles of information reacting to the input of the recommender system—which to some extent ignores the role of users as autonomous citizens or agents that should have the opportunity to control their information diets and selection options. Yet research in the social sciences indicates that users face specific ritualized, technical,

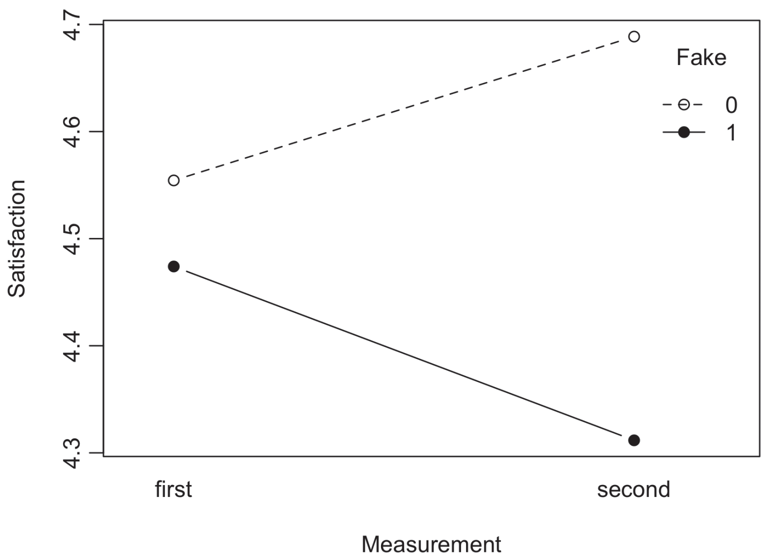

FIGURE 8.2 Interaction of measurement, control, and satisfaction.

and motivational barriers that hinder them from playing an active role in shaping a news recommender system so that it reflects their needs and wishes (Monzer et al., 2020; Powers, 2017). Moreover, customizable interfaces ask investment of time and energy of the respondent—something that does not necessarily go hand in hand with an easy-to-use system and enhanced user satisfaction. Actively customizing a recommender is a clear path towards agency; however, the next question is whether or not the effects of the customization can even be perceived, given the complexity of algorithmic recommendation systems.

In this study, we find that providing the functionality of explicit control of a news recommender system to users indeed leads to higher levels of perceived control. Yet this does not necessarily translate into high levels of usage of explicit control settings. On the contrary, in our study, participants barely engaged with the control panels, and if they did the engagement was not sustained. This raises questions for future research: first, what are the reasons for the lack of user interest in controlling a news recommender system? Second, does actively exercising control add to their feeling of agency, or is merely the perception of agency already enough to feel empowered? Third, more qualitative research is needed to understand how users exercise agency in news consumption, i.e., via communicative relations within news consumption networks. This also raises questions about how this could translate into the design of digital news consumption infrastructures.

Moreover, higher levels of control also did not lead to higher levels of user satisfaction. Concretely, this means that the economic incentives to offer more control may be limited. If there are no economic incentives to provide more user control, the only option to ensure controllability is to demand it by law. This could point to a role for the law in demanding more meaningful control options if such options are not offered otherwise. Yet these control options would need to be carefully designed to be empowering and intuitive enough so that users would at least try them from time to time. Having said that, we also found that providing an illusion of control that is not matched with actual control decreases satisfaction with the recommender system significantly. So, while adding control options does not add to user satisfaction, being dishonest about control translates into a negative evaluation of the system in general.

In this study, we employed a novel research design in which we could observe and survey users interacting with a recommender system in which we systematically controlled the parameters of the system. The advantage of this system is that it overcomes issues of external validity in experimental research, yet it also limits the generalizability of our results. First, our recommender only recommended news items, whereas in the reality of platform-afforded communication, news is mixed with private information and ads. Second, the limited sample size did not allow us to test one of the most commonly used algorithms for news selection, which is collaborative filtering. This algorithm matches users with other users that have engaged with similar content and recommends items these users also engaged with; to realistically simulate collaborative filtering we would have needed to include thousands of data points which is the beyond the scope of our experiment. Third, we did not include individual-level factors such as motivation to exercise control or digital literacy, to understand whether the willingness to exercise control and its effects are conditional on user types. Future studies should also test more and different types of control options to the users, in particular, options that are more intuitive and adaptive to the user. It could be that the limited enthusiasm of exercising control has been a function of the way control was designed and operationalized.

Nevertheless, our study provides valuable first insights into the conditions under which users exercise control in the context of a complex AI system. How far are humans really willing and able to use the opportunities the technology provides them to become the human in the loop in the process of news production, dissemination, and consumption? Our findings suggest that realizing the communicative role of technology (Guzman, 2019) is a complex endeavor. Possibly, the conceptualization of the user as a static object by those building recommender systems (Dahlgren, 2018) has led to an acceptance of that very role by users themselves—an acceptance that cannot simply be fixed by a normative need for users to be more than that, even if it is a crucial precondition for functioning democracies in a datafied society.

From a more optimistic point of view, it should be noted that extant work has established that users are in principle motivated to play a more active role (Monzer et al., 2020). However, this role is a departure from the delegation of trust towards journalistic gatekeepers that media users were socialized into. Thus, to realize user agency vis-à-vis recommender systems—and probably all AI systems—norms and practices of self-determined political communication as related to building political knowledge and collecting information need to be fostered in new ways (Hintz et al., 2019). This process is likely to take time. Yet, designing technology with this goal in mind could accelerate our capacities to understand, reflect on, and critically use technology, in order to feel like we are in control of our political information use in today's datafied society.

Notes

1 All hypotheses and analyses have been preregistered. The report can be found at https://aspredicted.org/t7hw5.pdf
2 The similarity was determined by using Soft Cosine Similarity (Sidorov et al., 2014).

References

Beam, M. A., & Kosicki, G. M. (2014). Personalized news portals: Filtering systems and increased news exposure. *Journalism & Mass Communication Quarterly*, *91*(1), 59–77. https://doi.org/10.1177/1077699013514411
Bodenhausen, G. V., & Hugenberg, K. (2009). Attention, perception, and social cognition. In F. Strack & J. Förster (Eds.), *Social cognition: The basis of human interaction* (pp. 1–22). Psychology Press.
Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, *15*(3), 209–227.
Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary effects of Facebook algorithms. *Information Communication and Society*, *20*(1), 30–44.
Chen, G. M., Chock, T. M., Gozigian, H., Rogers, R., Sen, A., Schweisberger, V. N., … & Wang, Y. (2011). Personalizing news websites attracts young readers. *Newspaper Research Journal*, *32*(4), 22–38. https://doi.org/10.1177/073953291103200403
Chung, D. S. (2008). Interactive features of online newspapers: Identifying patterns and predicting use of engaged readers. *Journal of Computer-Mediated Communication*, *13*(3), 658–679.
Cotter, K., & Reisdorf, B. C. (2020). Algorithmic knowledge gaps: A new dimension of (digital) inequality. *International Journal of Communication*, *14*, 745–765.
Dahlgren, P. (2018). Media, knowledge and trust: The deepening epistemic crisis of democracy, *Javnost – The Public*, *25*(1–2), 20–27
Gertler, B. (2017). Self-knowledge. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2017 ed.). Metaphysics Research Lab, Stanford University.
Go, E., Jung, E. H., & Wu, M. (2014). The effects of source cues on online news perception. *Computers in Human Behavior*, *38*, 358–367. https://doi.org/10.1016/j.chb.2014.05.044
Goossen, F., IJntema, W., Frasincar, F., Hogenboom, F., & Kaymak, U. (2011). News personalization using the CF-IDF semantic recommender. *Proceedings of the International Conference on Web Intelligence, Mining and Semantics*. https://doi.org/10.1145/1988688.1988701
Guzman, A. L. (2019). Prioritizing the audience's view of automation in journalism. *Digital Journalism*, 7(8), 1185–1190.

He, C., Parra, D., & Verbert, K. (2016). Interactive recommender systems: A survey of the state of the art and future research challenges and opportunities. *Expert Systems with Applications*, *56*, 9–27.
Helberger, N. (2019). On the democratic role of news recommenders. *Digital Journalism*, 7(8), 993–1012.
Hildebrandt, M. (2019). Privacy as protection of the incomputable self: From agnostic to agonistic machine learning. *Theoretical Inquiries in Law*, *19*(1), 83–121.
Hintz, A., Dencik, L., & Wahl-Jorgensen, K. (2019). *Digital citizenship in a datafied society*. Polity Press.
Jugovac, M., & Jannach, D. (2017). Interacting with recommenders – overview and research directions. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 7(3), 1–46. https://doi.org/10.1145/3001837
Kang, H., & Sundar, S. S. (2013). Depleted egos and affirmed selves: The two faces of customization. *Computers in Human Behavior*, *29*(6), 2273–2280.
Kannengießer, S. (2020). Engaging with and reflecting on the materiality of digital media technologies: Repair and fair production. *New Media & Society*, *22*(1), 123–139.
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
Knijnenburg, B. P., Reijmer, N. J., & Willemsen, M. C. (2011, October). Each to his own: How different users call for different interaction methods in recommender systems. Proceedings of the Fifth ACM Conference on Recommender Systems (pp. 141–148). https://doi.org/10.1145/2043932.2043960
Ku, O., Hou, C.-C., & Chen, S.Y. (2016). Incorporating customization and personalization into game-based learning: A cognitive style perspective. *Computers in Human Behavior*, *65*, 359–368.
Lee, A., & Delli Carpini, M. X. (2010). News consumption revisited: Examining the power of habits in the 21st century. Paper presented at the 11th International Symposium on Online Journalism, Austin, Texas, April 23–24, 2010. https://isoj.org/wp-content/uploads/2018/01/LeeCarpini10-1.pdf.
Livingstone, S. (2000). Television and the active audience. In D. Fleming (Ed.), *Formations: A 21st century media studies textbook* (pp. 175–195). Manchester 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. https://doi.org/10.5117/ccr2020.1.003.loec
McNee, S. M., Lam, S. K., Konstan, J. A., & Riedl, J. (2003). Interfaces for eliciting new user preferences in recommender systems. *Lecture Notes in Computer Science*, 178–187.
Min, S. J. (2019). From algorithmic disengagement to algorithmic activism: Charting social media users' responses to news filtering algorithms. *Telematics and Informatics*, *43*, 101251. https://doi.org/10.1016/j.tele.2019.101251
Monzer, C., Moeller, J., Helberger, N., & Eskens, S. (2020). User perspectives on the news personalisation process: Agency, trust and utility as building blocks. *Digital Journalism*, *8*(9), 1142–1162.
Morais Carvalho, J., Arga e Lima, F., & Farinha, M. (2021). Introduction to the Digital Services Act, content moderation and consumer protection. *Revista de Direito e Tecnologia*, *3*(1), 71–104.
Nguyen, T. H., & Nguyen, V. A. (2016). Implicit feedback mechanism to manage user profile applied in Vietnamese news recommender system. *International Journal of*

Computer and Communication Engineering, *5*(4), 276–285. https://doi.org/10.17706/ijcce.2016.5.4.276-285

Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin. https://doi.org/10.3139/9783446431164

Phelan, O., McCarthy, K., Smyth B. (2010). Buzzer – online real-time topical news article and source recommender. In L. Coyle & J. Freyne (Eds.), *Artificial intelligence and cognitive science: Lecture notes in computer science*, vol 6206 (pp. 251–261). Springer.

Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences*, *91*(16), 7398–7403.

Powers, E. (2017). My news feed is filtered?: Awareness of news personalization among college students. *Digital Journalism*, *5*(10), 1315–1335. https://doi.org/10.1080/21670811.2017.1286943

Russell, A., & Waisbord, S .(2017). The Snowden revelations and the networked fourth estate. *International Journal of Communication*, *11*, 858–878, https://ijoc.org/index.php/ijoc/article/view/5526

Sidorov, G., Gelbukh, A., Gómez-Adorno, H., & Pinto, D. (2014). Soft similarity and soft cosine measure: Similarity of features in vector space model. *Computación y Sistemas*, *18*(3), 491–504. https://doi.org/10.13053/cys-18-3-2043

Strycharz, J., Ausloos, J., & Helberger, N. (2020). Data protection or data frustration? Individual perceptions and attitudes towards the GDPR. *European Data Protection Law Review*, *6*(3), 407–421

Sundar, S. S. (2008). Self as source: Agency and customization in interactive media. In E. A. Konijn, S. Utz, M. Tanis, & S. B. Barnes (Eds.). *Mediated interpersonal communication* (pp. 58–74). Routledge.

Thorson, K., Xu, Y., & Edgerly, S. (2018). Political inequalities start at home: Parents, children, and the socialization of civic infrastructure online. *Political Communication*, *35*(2), 178–195.

Thurman, N., Moeller, J., Helberger, N., & Trilling, D. (2019). My friends, editors, algorithms, and I: Examining audience attitudes to news selection. *Digital Journalism*, 7(4), 447–469.

Zhu, D., & Lee, S. (2019). Autonomous readers: The impact of news customisation on audiences' psychological and behavioural outcomes. *Communication Research and Practice*, *6*(2), 125–142. https://doi.org/10.1080/22041451.2019.1644586

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.