

“I always assumed that I wasn’t really that close to [her]”: Reasoning about invisible algorithms in the news feed

ABSTRACT

Our daily digital life is full of algorithmically selected content such as social media feeds, recommendations and personalized search results. These algorithms have great power to shape users’ experiences yet users are often unaware of their presence. Whether it is useful to give users insight into these algorithms’ existence or functionality and how such insight might affect their experience are open questions. To address them, we conducted a user study with 40 Facebook users to examine their perceptions of the Facebook News Feed curation algorithm. Surprisingly, more than half of the participants (62.5%) were not aware of the News Feed algorithm at all. Initial reactions for these previously unaware participants were surprise and anger. We developed a system, *FeedVis*, to reveal to users the difference between the algorithmically curated and an unadulterated News Feed, and used it to study how users perceive this difference. Participants were most upset when close friends and family were not shown—they had often inferred social meaning from the filtering of the feed. By the end of the study, however, participants were mostly satisfied with the content on their feeds. Following up with participants two to six months after the study, we found that for most, satisfaction levels remained similar before and after becoming aware of the algorithm, however, algorithmic awareness led users to more actively engage with Facebook and bolstered their overall feelings of control on the site.

Author Keywords

Algorithms; Invisibility; News Feeds

INTRODUCTION

Today, algorithms curate everyday online content by prioritizing, classifying, associating, and filtering information. And in doing so, they exert power to shape the users’ experience and even their perception of the world [9]. News Feeds, which provide users with frequently updated news, are one application where algorithms play an influential role. For instance, while news of the protests in Ferguson, Missouri, USA dominated Twitter in August 2014, this was not the case on Facebook. A random sample of 100,000 U.S. Facebook users from the 9th to 20th of August showed that users were talking about the ‘*ALS ice bucket challenge*’ more than twice as much as the protests [26]. In investigating the reason for this difference, it was found that the Facebook’s News Feed ranking algorithm prioritizes stories posted by a user’s friends to

make them more relevant. However, a Twitter user sees all tweets of users she follows. So some argue that Facebook users might be isolated in a “filter bubble” [32], seeing information that Facebook thinks they want to see rather than what they might need to see.

While such powerful algorithms are omnipresent online, they are rarely highlighted in the interface, leaving users unaware of their presence. Although the lack of users’ awareness about these hidden processes can sometimes indicate a successful design, in some cases this invisibility can cause problems. A clear example is Morris’s study of social network use by new mothers. She questioned the common complaint that new mothers exclusively posted photos of their babies. She found that Facebook News Feed created this misperception because it prioritizes posts that receive likes and comments – photos of babies often received attention from a large audience. Because users lack knowledge about the News Feed algorithm, they may have an inaccurate picture of how their and others’ actions influence their personal feeds [29].

The increasing prevalence of opaque and invisible algorithms coupled with their power raises questions about how knowledgeable users are and should be about the existence and operation of these algorithms. Whether their understanding is correct or not, users’ perceived knowledge about an algorithm can affect their behavior. For instance, believing that posts with commercial keywords were ranked higher by the Facebook News Feed algorithm led some teenagers to add product names to their posts in an attempt to manipulate the algorithm and increase their posts’ visibility [6].

However, with no way to know if their knowledge of these invisible algorithms is correct, users cannot be sure of the results of their actions. While this indicates that increased knowledge may result in behavioral changes for some, it remains an open question whether it is useful to give users some insight into algorithms’ existence or functionality, in general. Beyond whether it is useful, we might also ask how this insight will affect their future interaction experiences. Particularly in social media, the opacity of these algorithms may affect users’ knowledge and social interactions in potentially negative ways, leading them to different understandings of current events or convincing them to block new mothers [40].

To begin to address these questions, we explored users’ awareness and perception of the Facebook News Feed curation algorithm (hereafter “*the algorithm*”). This algorithm determines which stories (e.g. status updates, pictures, videos, likes and comments) appear in a Facebook user’s News Feed based on social network links and activity on Facebook [19, 18]. We interviewed 40 Facebook users and discovered that more than half (62.5%) were not aware that News Feed hid stories. They believed every single story from their friends and followed pages appeared in their News feed. To under-

Paste the appropriate copyright statement here. ACM now supports three different copyright statements:

- ACM copyright: ACM holds the copyright on the work. This is the historical approach.
- License: The author(s) retain copyright, but ACM receives an exclusive publication license.
- Open Access: The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.

This text field is large enough to hold the appropriate release statement assuming it is single spaced.

stand why so few participants knew of the algorithm's existence, we investigated their Facebook usage patterns, finding associations between awareness and Facebook usage.

We developed *FeedVis*, a Facebook application, to reveal the algorithm to study participants. *FeedVis* extracted participants' News Feed stories as well as their friends' stories to disclose what we call "*the algorithm outputs*": the differences in users' News Feeds when they have been curated by the algorithm and when they have not. Using *FeedVis*, we showed participants alternate views of their familiar News Feed to understand how they reacted to the algorithm outputs. We finally provided them with an opportunity to modify the algorithm outputs to curate their desired News Feed. We discovered that strong initial reactions often subsided once users understood who and what was being hidden. We followed up with participants two to six months later and found that their usage behaviors had often changed due to the insight they gained about the algorithm via our study.

RELATED WORK

Many areas of research have examined invisible processes and how people react to them. Cognitive science and human factors researchers study the mental models people create and cognitive structures that develop when interacting with hidden processes of machines and technology [30]. To find new design ideas, designers have proposed probes of people's interactions with hidden and uncertain aspects of their lives [12]. Related efforts exist in architecture and urban planning, studying how people perceive and navigate urban landscapes. This work helps designers to gain insight for good urban design [25]. Finally, time and motion studies observe people conducting a task and extract any hidden patterns to find the most productive way to complete it [14]. Studies dealing with hidden or invisible components of daily life have also addressed some aspects of social media. The invisibility of audiences in online environments has prompted research into the imagined audience [24], including quantifying how perceived audiences compare to actual audiences and measuring invisible currents of attention on social media [4].

Algorithms

Algorithms, as invisible and influential pieces of daily digital life, have become the focus of research attention. Many researchers have looked at particular types of algorithms and argued that their effects are important but their operations are opaque [2, 17, 35]. For example, search algorithms may structure the scope of online information access for society, functioning as gatekeepers and creating a politics of search [16, 20]. Targeting of ads has been studied by researchers, arguing that the opacity of the algorithms may mask bias or discrimination in the results, such as the uneven distribution of arrest record ads by race [39]. And ranking of journal articles has been found to potentially result in unintended differences in the perceived importance of scientific articles [7].

Researchers have paid particular attention to algorithms when outputs are unexpected or when the risk exists that the algorithm might promote antisocial political, economic, geographic, racial, or other discrimination. Invisible algorithms in health care, credit scoring and stock trading have aroused

interest in recent years [33, 38]. Other researchers have looked at dynamic pricing and the possibility of reinforcing biases against rural and poorer areas, which tend to have less competition, thereby "diminish[ing] the Internet's role as an equalizer" [41]. Controversy over Twitter Trends and accusations of algorithmic censorship of the term #occupywallstreet throughout the Occupy Wall Street protests led to questions of whether a sorting algorithm can be wrong or unethical under some conditions [15]. Some researchers have even studied unexpected results in the filtering of autocompletion text, finding some algorithms explicitly attempt to make moral judgements, like removing terms for child pornography [8].

As a result of these concerns, some have argued that increased algorithmic transparency would be beneficial. Designs and recommendations have been developed to reveal the power of algorithms to predict people's interests and affecting their online life [11, 28]. For example, the campaign '*digital shadow*' accesses Facebook users' profiles with their permission to show them how much personal information is available for algorithms to use and how much it is worth [10].

Algorithmically Generated Feeds

The prevalence of algorithmically generated feeds in social media such as the Facebook News Feed and the Twitter Feed has triggered discussions about the appropriateness of the curation algorithms employed. For example, some have argued that filtering and prioritizing stories might make some friends vanish [34]. This vanishing effect mainly impacts business owners who use social media as an advertising channel [5]. The profit incentive leads advertisers to find ways to keep a story near the top of feeds longer. One of the primary methods used is reverse engineering the feed curation algorithms to understand how they work. Reverse engineering of algorithms is sometimes used even by regular users to ensure their stories are shown on others' feeds [17].

While tools have been developed to show summaries of algorithmic results, to our knowledge no researchers have developed systems to reveal to users the contrast between algorithmically manipulated and neutral results. In our work, we develop such a tool and use the Facebook News Feed curation algorithm as an example. Launched in 2006 [36], the Facebook News Feed curation algorithm has attracted significant attention in recent years, particularly after a recent, controversial study of emotional contagion [23]. Facebook currently uses close to 100,000 factors to choose the best stories from the large pool of potential stories [27]. Although Facebook has stated it would change how it communicates updates to the News Feed due to the large number of user requests [1], there is still little understanding among users or anyone outside of Facebook of how the News Feed curation algorithm works. To shed light on invisible algorithms curating social media feeds and how they impact users, we ask the following research questions:

RQ1. How aware are users of News Feed algorithmic manipulation and what factors are associated with this awareness?

RQ2. How do users react to their News Feeds' curation when shown the algorithm outputs? Given the opportunity to

personally alter the outputs, how do users' preferred outputs compare to the algorithm's?

RQ3. How does the knowledge users gain through an algorithm probe tool transfer to their usage behavior?

STUDY DESIGN

In order to address the proposed research questions, we conducted a mixed-methods study consisting of three phases. First, participants visited our laboratory and completed a questionnaire and interview to measure algorithm awareness. At this time, we also collected participants' network size, NewsFeed stories and friends' stories to populate an interface for the next phase. Second, during the same visit, participants used an application to visualize the algorithm outputs, and we used a long form open-ended interview to discuss them. Third, we e-mailed participants two-six months later to ask closed- and open-ended questions to evaluate the consequences of any insight gained by observing the algorithm outputs. All in-person interviews were audio recorded and transcribed for analysis.

Pre-Assessment: Testing Algorithm Awareness

In the beginning of the study, the participants answered a demographic questionnaire including measures of their social media use. To assess their familiarity with the algorithm, we asked a combination of open- and closed-ended behavioral, knowledge, and attitude questions whose answers likely depend upon their awareness of the algorithm. First, we asked if and how they used Facebook settings to adjust the content on their News Feed (including sorting the stories of News Feed based by recency or top stories, hiding a story, following or unfollowing friends and making Facebook lists). Next, we asked them to imagine they had a 'friend,' Sarah, and she shared a public story visible on her wall to all her friends. We asked them whether this story would appear in their own News Feed. In addition, we asked whether they missed any stories that they would have preferred to see in their News Feed. If they answered affirmatively, we probed further to understand their reasoning for why they may have missed a story; for instance, whether they thought missing a story would be a result of their own actions such as scrolling past it quickly or they considered the existence of a filtering process as a possible reason. During this pre-assessment, we asked participants to use their Facebook accounts to log into our Facebook application, FeedVis. FeedVis then extracted and collected the participant's network size, News Feed and their friends' public stories. The News Feed and friends' stories were used to generate a series of alternate views for the feed; the network size was used to explore associations with algorithm awareness.

Main Interview: Algorithm Outputs Disclosure

After understanding the participants' existing knowledge, we presented a series of FeedVis NewsFeed views. Paging through these views revealed some algorithm outputs to the participants and made them aware of the News Feed's algorithmic curation, if they were not aware already. As extracting all stories from all friends is processing-intensive, we limited the time period of the stories collected to one week or less depending on the number of the user's friends. We described

four FeedVis views of their feed to the participants: The Content View, Friend View, Friend Rearrangement and Content Rearrangement Views.

Feedvis Content View: Revealing Content Filtering

The content that the Facebook algorithm shows a user is chosen from the universe of all stories contributed by the people and pages that a user follows. In the first view, we aimed to show the user this universe of potential content, highlighting content that was not chosen for display. This view helped the user compare what they saw and what they might have seen in the absence of a filter, or with a different one. The Content View consisted of two columns (Figure 1). The right column, 'Shown stories,' displayed only the stories shown on the user's News Feed. These stories were shown with a blue background. The left column, called 'All stories,' showed every story from all the users' friends and all the pages the user follows. In this column, stories which did appear in the user's News Feed were again shown on a blue background, while stories which did not appear in their News Feed were shown on a white background. To create this column, we extracted all viewable stories from the page of each of the participant's friends and checked the information provided by the Facebook API to determine whether it appeared in our participant's News Feed or not.



Figure 1. Content View. Shown stories (in blue) occur across both columns, while the hidden stories (in white) appear only in the left column as 'holes' in News Feed.

Feedvis Friend View: Revealing Social Patterns

In addition to affecting the display of content, the Facebook algorithm could affect how a participant's friends are perceived by altering the frequency with which they appear in the feed. We built a visualization, Friend View, to help the user understand whose stories appeared and whose were hidden in their News Feed. This view divided the users friends into three categories based on the proportion of each friends stories that had appeared in the users News Feed during the previous week: 'rarely shown', 'sometimes shown', and 'mostly shown' friends (Figure 2).

Feedvis Friend & Friend Rearrangement Views: Envisioning a Different Algorithm

After exploring the algorithm outputs, we wanted to gauge participants' desire to change those outputs. Therefore, we created two new views that invited participants to "tweak" their algorithm. One allowed for adjustment based on authorship of stories, and the other invited manual filtering based

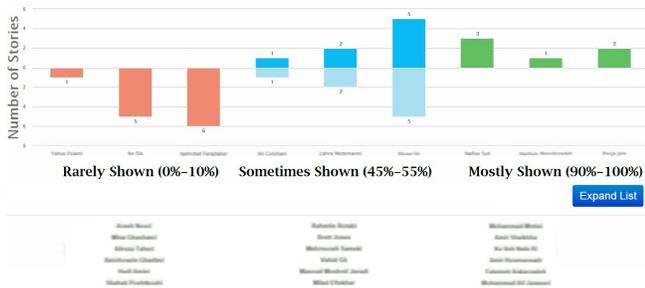


Figure 2. Friend View. ‘Rarely shown’ includes friends whose stories were mostly hidden (0%-10%) from the user. ‘Sometimes shown’ includes friends who had around half of their posts (45%-55%) shown to the user. ‘Mostly shown’ includes those friends whose stories were almost never filtered out (90%-100%) for the user. The number of the shown stories are displayed above the x-axis and the number of hidden stories are below the x-axis. Expand button adds more friends to each category, shown under the chart.



Figure 3. Friend Rearrangement View. User can move friends between the categories by changing the color of a friend to the destination category’s color.

on the content of stories. The former, Friend Rearrangement View (Figure 3), presented a list of friends according to the same three categories described above, and invited re-assignment of friends to different categories. The latter, Content Rearrangement view (Figure 4), selected ten seen posts and ten unseen posts at random, inviting users to indicate for each post whether they would have preferred a more or less restrictive filter. The lab portion of this study, including the pre-assessment, lasted from one to three hours per participant.

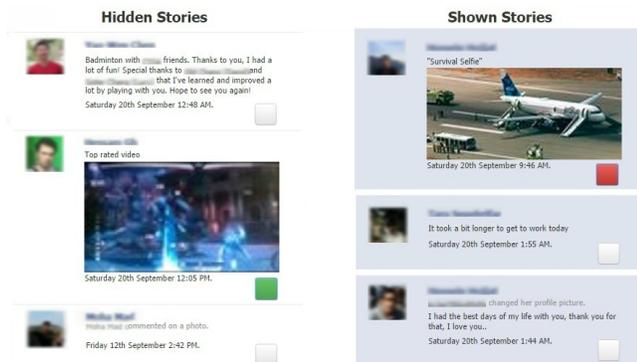


Figure 4. Content Rearrangement View. User can move a story from its original category to the other by clicking the button beside each story.

Post-Assessment: Evaluating Algorithm Outputs Revelation

To understand the impact of revealing the hidden aspects of the algorithm, we contacted participants via e-mail two months after conducting the study. First, we asked them whether participation in the study resulted in more, less or no change in their satisfaction with the Facebook News Feed. Then we asked them whether and how they changed their Facebook usage behavior after visiting the lab.

Participants

We used modified quota sampling to obtain a non-probability sample that is roughly representative of the US population on four dimensions. The national proportions for gender, age, race/ethnicity and socioeconomic status were used as quota targets for recruitment and selection in a Midwestern city. Quotas required an elaborate recruitment strategy including posters in varied public places, e-mails to local online communities and civic organizations, and posts on Facebook. We recruited 40 participants consisting of five students, two faculty members and 14 staff from a large Midwestern university and 19 people with other occupations such as homemakers, delivery persons, servers, bartenders, artisans, performers and writers. Participants received \$10/hour for the pre-assessment and main interview; participation in the post-assessment entered them in a lottery for a \$50 gift card. The original sample was 60% women and ranged between 18 and 64 years old. 68% of the participants were Caucasian, 15% were Asian and the African-American, Hispanic and Native American participants were nearly equally distributed. Approximately half of the participants’ annual income was less than \$50,000 and the rest were between \$50,000 and \$150,000. Our participants are typical of Facebook users in terms of age, gender, race and income [37, 3].

Data Analysis

To organize and conceptualize the main themes discussed by the participants, two researchers used line-by-line open coding to label the pre-assessment, main interview, and post-assessment data under primary categories and subcategories. We used Nvivo [31] to map the interviewees’ statements to these categories. Through a collaborative, iterative process, we revised these categories to agreement, then used axial coding to extract the relationships between themes. To further explore our data, we used statistical analysis to support our qualitative findings. For clarity, details of this analysis will be presented later in the paper.

RESULTS

Awareness of the Algorithm (RQ1)

Surprisingly, the majority of the participants (62.5%) were not aware of the algorithm’s existence. When asked whether the public story of their ‘friend’, Sarah, would definitely be shown in their News Feed, they answered affirmatively: *“I bet it would be on my News Feed. I probably would catch [it] at some point during the day.”* (P30). In their opinion, missing a public story was due to their own actions, rather than to those of Facebook. Importantly, these participants felt that they missed friends’ stories because they were not observing News Feed constantly or carefully. This was either by scrolling the News Feed too quickly or visiting Facebook too infrequently. They believed if they *“wanna go back to [a*

missed story], it's accessible" (P39) in their News Feed. We refer to this majority as the 'Unaware' participants.

The rest of the participants (37.5%) knew that their News Feed was filtered. When answering the question about Sarah's story, they stated that a friend's story might not appear in their News Feed due to a filtering process: "I don't think everything is supposed to be there. I mean I don't think the News Feed shows everything that everyone puts on Facebook. It's just certain things." (P22). As a result of their knowledge, these participants stated that they might miss a story because of the Facebook algorithm in addition to their own actions. We refer to them as the 'Aware' participants.

Paths to Awareness

We investigated Aware participants' responses further to understand how they became aware of their News Feed manipulation when so many others did not. Three participants learned of the algorithm's existence from external sources such as other people and news articles. However, most Aware participants stated they gained knowledge about the algorithm themselves via one of two common paths: inductively comparing feeds vs. deductively considering network size.

Inductively Comparing Feeds: Most Aware participants (n=12) mentioned that they compared which friends' stories appeared in their News Feed with other users. Noticing the discrepancy between the number of displayed stories from those friends, they felt they were seeing some friends' stories much more than the others. This observed difference suggested to them the possibility of the existence of a News Feed filtering process: "I have like 900 and some friends and I feel like I only see 30 of them in my News Feed. So I know that there's something going on, I just don't know what it is exactly." (P26). We asked these participants for greater detail to understand how and why they had noticed these differences. Most had observed that interacting with a friend (e.g. visiting their page, liking and commenting on their stories) often resulted in more stories from that friend in their News Feed. A few compared their News Feed to their friends' pages and found that stories were missing.

Deductively Considering Network Size: Seven Aware participants believed a filtering process must logically be part of the News Feed curation by necessity, since "there's too much material in general on Facebook." (P22). They argued that as the number of friends that people have on Facebook increases, there should be "some way that filters out those [stories] that you may not be as interested in." (P31). These participants thought of the algorithm as a basic and even obvious element in curating News Feeds that must exist in order to avoid overwhelming readers.

Although there were many avenues towards algorithm awareness, more than half of the participants were unaware of the algorithm's existence. This raises questions about their unawareness: While all the participants were exposed to the algorithm outputs, why were the majority not aware of the algorithm? Were there any differences in Facebook usage associated with being aware or unaware of News Feed manipulation? The following section answers these questions.

Connecting Exposure and Engagement to Awareness

To address the above questions, we investigated the participants' Facebook usage behavior. Some participants engaged with the algorithm outputs provided for them *passively* by, for instance, scrolling the News Feed and reading the stories as they appeared. On the other hand, some participants engaged with the algorithm outputs *actively* by, for example, adjusting their News Feed content using the settings Facebook provided. To understand whether this difference in engagement with the algorithm outputs was associated with algorithm awareness and to identify features related to these engagement patterns, we turned to a combination of our interview data and the data we extracted from each participant's Facebook. We identified three passive and four active engagement features. Each of these features were either mentioned by the participants or found in their Facebook data.

Passive Engagement: We identified several features that are likely to be related to awareness of the algorithm, but that may not imply any intentional activity by the user or could involve circumstances that are out of their control. These include: *Membership duration*; the number of years a user has been a member of Facebook. *Shown content percentage*; the ratio of the number of stories in a user's News Feed to the number of all the potential stories that could have appeared in an unfiltered News Feed. A smaller shown content percentage means overall the user would expect to read fewer stories from any friend. *Friendship network size*; the number of Facebook friends. Network size can be grown actively or passively — for example by responding to friend requests initiated by others — and it may reflect social behavior outside of Facebook (such as actual friendships) rather than decisions related to the platform. Network size is related to algorithm awareness because the limited space in the News Feed causes a greater proportion of potential stories to be filtered by the algorithm when the network is large¹.

Active Engagement: We then identified several features that are related to awareness of the algorithm and are more likely to also indicate platform- or algorithm-related intentional behavior. They are: *Usage frequency*, the number of times per day a participant uses Facebook. Frequent users are more prone to active engagement with the algorithm outputs. They exploring more spaces on Facebook (such as options and settings screens) and are more likely to comparing different aspects of feeds with each other. *Activity level*, a categorization of users as listeners (mostly reading News Feed without posting a story), light posters (posting stories occasionally), or heavy posters (posting stories frequently) based on survey and interview replies. A light or heavy poster is more actively engaged with algorithm outcomes than a listener because they receive feedback and attention (likes and comments) to their stories and this affects the algorithm's behavior. This makes a potential filtering process more salient. *News Feed content adjustment*, whether a participant uses settings to control what they see in their News Feed. Sorting stories based on the importance, following a friend, hiding a story and making lists are some examples of these settings. Using each of these options make a user more actively engaged with the algorithm outputs because they are trying to change those out-

¹We found friendship network size and shown content percentage have a significant negative correlation; $r = -0.44$, $p = 0.005$

comes. *Facebook page/group management*, whether a user is involved in managing a Facebook page or group. This suggests familiarity with Facebook analytics (information that shows a page manager how many people see a page’s story, revealing the existence of a filtering process).

We used open coding to find and compare engagement patterns between the Aware and Unaware participants using these features. We then used statistical methods to support our qualitative analysis. For numerical features, we conducted Welch’s test to avoid unequal sample size and variance side effects between the Aware and Unaware groups. For categorical features, we used Chi-square tests. We then ran Fisher’s exact test to confirm the Chi-square results and avoid possible effects due to small sample size.

We found a significant difference between the Aware and Unaware groups for all of the active engagement features by both thematic and statistical analysis (Table 1). In terms of usage frequency, we found that all the participants who had high usage frequency (more than 20 times in a day) were aware of algorithmic manipulation. Statistical analysis supported this finding by showing significant difference in usage frequency between the Aware (M=27.18, SD=33.8) and Unaware participants (M=6.92, SD=5.79). These frequent users used Facebook “*all day*” (P21), they were “*constantly logged in*” (P33) and looked at Facebook “*too many [times] to count*” (P22). We hypothesize that spending more time on Facebook let these participants explore more stories, features and spaces (such as the News Feed and others’ profile pages) than infrequent users. This exploration led to inductive feed comparisons and consequently new knowledge about News Feeds and the algorithm.

Table 1. Active Engagement Features

Active Engagement	t-value [§] Chi-square [†]	p-value	Effect size
Usage Frequency	-2.2 [§]	0.03	0.83
Activity Level	8.57 [†]	0.03	0.46
News Feed Content Adjustment	14.14 [†]	0.00	0.59
Facebook Page/Group Management	4.23 [†]	0.04	0.32

Unlike Unaware participants who labeled themselves posters or listeners, all 15 Aware participants declared themselves posters (light or heavy). In Aware participants’ discussions of their Facebook usage, we found that the number of likes and comments on their own stories suggested the possibility of the existence of a filtering process. These participants found that their popular stories were shown in their friends’ News Feeds more often. “*So I feel some of the stuff got to reach to [a] certain threshold of comments or number of likes before Facebook thinks that I might be interested in [it]*” (P23)

All six participants who did not apply any settings to adjust their News Feed content were unaware of algorithmic manipulation of their News Feed. Conversely, all the Aware participants tried to adjust their News Feed content by using at least one of the options provided by Facebook. Among the participants who did not apply any changes to their News Feed, some believed they “*cannot control the News Feed [since] it’s kind of receiving what Facebook gives [us], it’s kind of limited.*” (P1). The rest in this group believed they could apply any settings to adjust their News Feed if they were “*willing*

to invest the kind of time to find out how to do them.” (P3), but they did not invest this time.

There were seven participants involved in Facebook page/group management and all were aware of News Feed manipulation. These participants mentioned that Facebook provided some analytics for page/group managers such as ‘post reach’ (the number of people in whose News Feed a page/group stories appeared) and ‘people engaged’ (the number of people who have clicked, liked, commented on or shared a story). They stated that observing this analytic information suggested a filtering process that causes some of their page/group stories to reach more people than the others: “*[My friends] all don’t get to see everything, and I’ve always been suspicious of [Facebook], on how they choose who gets to see it, who doesn’t.*” (P28). Consistent with theories about the construction of mental models [21, 22], we believe these participants extended their knowledge from a known domain (Facebook page/group) into an unknown domain (personal profile) and used the analogy between these two domains to infer the algorithm’s existence in their personal profiles.

In contrast to the active engagement features, we did not find any noticeable difference between the Aware and Unaware groups in terms of the passive engagement features. This suggests that being a periodic Facebook user over many years, having a larger friendship network, or having a smaller fraction of stories from your friends actually shown in your News Feed is not associated with an awareness of the algorithm. These results suggest that simple exposure to the algorithm output is not enough to gain information about the algorithm’s existence. To learn about an algorithm without any outside information, active engagement is required.

Reactions to & Expectations of Algorithm Outputs (RQ2)

Once we understood the participants’ prior awareness of the algorithm’s existence, we walked them through the FeedVis tool. We started with the Content and Friend Views, to discover how they would react to alternative algorithm outputs. Then we directed them to the Friend and Content Rearrangement Views, giving them the opportunity to create their desired Friend and Content Views.

Initial Reactions

Many of the Unaware participants (n=15) were initially very surprised by how long the ‘all stories’ column was in comparison to the ‘shown stories’ column in the Content View (Figure 1): “*So do they actually hide these things from me? Heeeeeey! I never knew that Facebook really hid something!*” (P1). One participant described it as a new concept that she had never considered before, despite using Facebook daily: “*It’s kind of intense, it’s kind of waking up in ‘the Matrix’ in a way. I mean you have what you think as your reality of like what they choose to show you. [...] So you think about how much, kind of, control they have...*” (P19).

Observing the algorithm outputs in FeedVis surprised some Unaware participants (n=11) by revealing misperceptions about their friends whose stories were not shown in the participants’ News Feed at all. For example, seven of them assumed that those friends simply did not post on Facebook. It

was through FeedVis that they discovered these friends did indeed post. A few participants falsely believed that those friends had left Facebook: *"I know she had some family issues so I just thought she deactivated her account."* (P35). Importantly, some participants disclosed that they had previously made inferences about their personal relationships based on the algorithm output in Facebook's default News Feed view. For instance, participants mistakenly believed that their friends intentionally chose not to show them stories because they were not close enough. Again they were surprised to learn via FeedVis that those hidden stories were removed by Facebook and not their friends: *"I have never seen her post anything!, and I always assumed that I wasn't really that close to that person so that's fine. What the hell?!"* (P3).

A few participants (n=5) were curious and began asking questions about the algorithm such as *"Do they choose what they think is the best for me to see? Based on what?"* (P37). This curiosity led them to start wondering whether *"there is some algorithm or something or some rules to choose these [hidden] things that would not appear [in News Feed]."* (P1). In contrast to Unaware participants, most of the Aware participants did not express surprise or curiosity, because of their previous awareness of the algorithm's existence. They did, however, express dissatisfaction as we describe below.

Expectations

Along with surprise and curiosity, many participants, Aware or Unaware, (n=19) expressed dissatisfaction and even anger when missing stories were revealed to them on FeedVis because Facebook violated their expectations: *"Well, I'm super frustrated [pointing to a friend's story], because I would actually like to see their posts."* (P3). Participants explained that seeing an otherwise hidden story would affect their behavior toward the friend who posted it: *"I think she needs support for that; if I saw it, then I would say something [to support her]."* (P8). In the Friend View, as with the Content View, many participants (n=19) expected their network to be categorized differently than the Facebook algorithm. This expectation was particularly high for family members, with many participants stating that family members should be in the 'mostly shown' category: *"I cannot really understand how they categorize these people. Actually this is my brother [in 'sometimes shown'] and actually he needs to be here [in 'mostly shown']."* (P1).

Along with such dissatisfaction, some participants (n=9) believed it was not Facebook's place to decide what to show in their News Feed: *"It was sort of like someone was deciding what I wanted to see and it kind of made me mad."* (P32). These participants preferred to see every single story and use *"manual filtering"* (P23) themselves. However, a few argued that Facebook, as a free service, had the authority to manipulate the feed without concern for the users' desires: *"I feel like I'm a mouse, a little experiment on us. To me, that's the price I pay to be part of this free thing. It's like we're a part of their experiment and I'm okay with it."* (P21).

Despite some of the frustration in their initial reactions and expectations, more than half of the participants (n=21) became more satisfied with the algorithm over the course of the study. Even as they first scrolled down the Content View,

many mentioned that they began to understand why Facebook hid some stories from them². For example, many hidden stories were about friends' interactions with each other (e.g. likes, comments, happy birthday messages) that were not relevant to them: *"A lot of what is filtered out are things that don't really pertain to me. I'm so grateful because, otherwise, it would just clutter up what I really want to see."* (P13).

To better understand how participants' expected outputs compared to the actual algorithm outputs, we first asked participants to move friends to their desired categories via the Friend Rearrangement View (Figure 3). On average, participants moved 43% of their friends to another category. This high rate of change demonstrates that the algorithm is not effectively capturing the strong feelings participants had about which friends' should appear in their News Feed. In the Content Rearrangement View (Figure 4), participants moved on average 17% of their News Feed content between the 'shown' and 'hidden' categories (SD = 9%), a noticeably lower percentage than the Friend Rearrangement View. Although many participants were initially shocked, concerned or dissatisfied with the existence of a filtering algorithm, they concluded that there were not many stories they actually wanted to move: *"Honestly I have nothing to change which I'm surprised! Because I came in like 'Ah, they're screwing it all!'"* (P23). These findings suggest that while filtering is both generally needed and appreciated, a lack of awareness of the existence of this process results in concern and dissatisfaction.

Transferring FeedVis Insight to News Feed (RQ3)

During our initial discussions with Aware participants, we found their perceptions about the algorithm already affected their Facebook usage. Awareness of the algorithm led them to actively manipulate their News Feed, using folk theories they developed about how the algorithm might work. For example, those who believed interacting with their friends would affect the number of stories seen from those friends adjusted their interactions: *"I know that if you don't interact with people you won't see their posts; sometimes I purposely don't interact with people just so that 'hahaha, manipulating the system'"* (P20). There were also participants who thought the number of stories displayed was limited by the algorithm. They believed if they unfollowed someone, *"there's always a new person that [would] start showing up more."* (P26).

In addition to manipulating their own News Feeds, a few Aware participants (n=4) tried to manipulate the News Feeds of others. Participants who believed that stories with more comments and likes would reach more people might comment on their own stories to get into more people's News Feeds. One suggested *"if you post a picture, without a comment, it's less likely to show up on your friends' News Feed."* (P21). This goal of affecting others' News Feeds was observed most among Aware participants who managed a Facebook business page. Since they saw their page followers as potential customers, they tried to increase their number of followers.

²As participants explored the algorithm outputs via the Content and Friend Views, we asked them to speak aloud, describing any patterns that might emerge. They described and revised fascinating folk theories explaining the algorithm. These theories are out of the scope of this paper and will be discussed in later work.

While one way to increase this number is buying fake likes, one participant argued that, due to the algorithm, this practice might decrease their profit: “[*Suppose that I’m going to buy more likes and all of the sudden I had 2000 more [...] But what happens is they have a whole bunch of fake people [...] So then if they’re sending [a story] out to 10% of the people and if you have 2500 likes, 250 of them are getting it. But if 90% of those are fake, then fewer real people are seeing it. So it doesn’t help you at all.*]” (P28).

Following Up with Participants

To understand whether exposure to the algorithm outputs during the study would prompt similar behaviors in the previously Unaware participants (or reinforce these behaviors among the Aware participants), we contacted our participants two to six months after the study. We asked them whether their Facebook usage or satisfaction with the News Feed had changed as a result of participating in our study. 30 out of 40 participants responded³.

Usage

Most (83%) reported changes in their behavior due to participation in our study. We noted that despite coming into the study with varying levels of awareness, both Aware and Unaware participants reported similar changes.

Manipulating the Manipulation: 21 out of the 30 who completed the follow-up (both Unaware and Aware) asserted that they started to manipulate what they saw on Facebook, mainly by using News Feed settings or changing their interaction with some friends. Of those who started to use News Feed settings for the first time after the study (n=13), most began using ‘most recent’ and ‘top stories’ options provided by Facebook to sort stories. Most said that they “*make more of an effort to make sure [their] viewing of posts is more on the ‘Most Recent’, as opposed to the ‘Top Stories’ option.*” (P35). A few stated that they “*tend to switch up between the ‘Most Recent’ setting and the ‘Top Stories’ setting.*” (P14).

10 participants changed their interaction with their friends in order to affect the stories appearing from those friends in their own News Feed. Some started to be “*more selective about clicking ‘like’ because it will have consequences on what [they] see/don’t see in the future.*” (P4). On the other hand, a few participants ‘liked’ more stories than they used to. This was particularly true if they “*may not want to comment on their status but want to make sure that their posts continue to show up in News Feed.*” (P31). A few participants changed their interaction with some friends by visiting their personal pages “*so they pop up on News Feed again.*” (P11). In addition, a few who realized that they might not see some of their friends’ stories due to the filtering process, said they were “*more likely to visit home pages for certain friends to see if they’ve posted anything.*” (P38). Finally, unfriending people in order to receive updates only from those they were most interested in was a more drastic change some participants mentioned.

³We attribute this attrition rate in part to the different incentives for participation in each part of the study. Initial lab visits were paid by the hour, while completing the e-mail follow-up entered participants into a lottery.

A few participants tried to make their own stories appear on more of their friends’ News Feeds. For example, they started to like their own posts “*to give them more visibility.*” (P28). Others modified their Facebook settings to limit which people received their stories.

Exploration: Four participants began to “*play around with Facebook a little more.*” (P25). They stated that after the study, they “*went back and started experimenting a little with the News Feed and discussing with some friends on ways to streamline*” (P10) what they were receiving in News Feed. Some also shared “*what learned from the study with others*” (P18) as they felt more knowledgeable about how Facebook worked. One participant even made their friends aware that the algorithm hid their stories from her News Feed: “*I told some friends that I was not seeing their posts.*” (P36).

Decreasing Usage Frequency: Three participants used Facebook less than they had in the past. One reason was the frequent changes to the News Feed settings, including the location of ‘Most Recent’ option, leaving them frustrated with the need to search for settings or understand their function. In an extreme case, one participant stopped using Facebook as she believed it was not straightforward with its users about curating the News Feed: “*After the study, I stopped using Facebook because I felt the way the Feed items were curated had, in some ways, broken the expectations between myself and Facebook [...] By neither showing me everything nor making their actions explicit, I felt like I was being lied to.*” (P3).

Overall, participation led to more informed Facebook use, even for those who were previously aware of the algorithm’s existence: “*It definitely made me more aware of how I was using it.*” (P20). Even the few participants who reported no change in their usage (n=6) noted they “*do feel more knowledgeable of the way [Facebook] ‘studies’ viewing preferences and accordingly adapts News Feed*” (P22) after the study.

Satisfaction

In the follow up, we also asked the participants whether participation in our study affected their satisfaction with News Feed. The majority of the participants (n=24) who answered this question reported the same or higher satisfaction level with News Feed after the study. However, a few participants (n=6) declared that their satisfaction decreased when they understood that “*some updates were deliberately not shown*” (P9). They explained that worrying they might miss stories they wanted to see made them trust News Feed less: “*I’m disappointed because I keep thinking that I might be missing some of the updates from my friends. [...] I don’t really trust the News Feed about giving me updates on everything I want to know.*” (P17). They also discussed that they felt “*less empowered to have an optimal experience [since] the rules can change at any time [...] which makes no promises in terms of permanence.*” (P21).

Participants who had the same or higher satisfaction level with News Feed generally discussed how they felt more knowledgeable about the algorithm as a result of participating. For instance, one Unaware participant stated that becoming aware of the algorithm’s existence resulted in less dissatisfaction when stories did not receive enough attention from

others: “Because I know now that not everything I post everyone else will see, I feel less snubbed when I make posts that get minimal or no response. It feels less personal” (P38). Another noted how understanding that Facebook hid some stories they might not be interested in made them “more interested in checking Facebook because it does not seem as cluttered with random information.” (P10). Overall, gaining insight about the algorithm via FeedVis caused people to say that they used Facebook more knowledgeably and their satisfaction level with Facebook generally remained high.

LIMITATIONS

We hope a larger sample of users will confirm the findings of this study, and it may be useful to increase the geographic diversity of users (a demographic we did not vary) to better match the US population. A limitation of our study also concerns our Facebook data. The Facebook API permits 600 queries per minute. As users with hundreds of friends or stories participated, our development decisions became query-limited. Furthermore, we sometimes struggled to retrieve reliable data from Facebook, as we occasionally observed incomplete results from the Facebook API.

DISCUSSION

This study indicates the importance of research into user experiences with algorithmically curated content in social media. While developers might expect that users experience social media with algorithms in mind, we found a very different reality. Most users were not aware of the existence of algorithmic curation despite using the Facebook News Feed an average of ten times per day. We found that users’ awareness of the filtering algorithm is prompted partly by certain kinds of active engagement with the algorithm outputs, like adjusting the News Feed content via Facebook settings.

On Facebook, ignorance about the algorithm had serious consequences. Our participants used the News Feed to make inferences about their relationships, wrongly attributing the algorithm’s actions to be the intent of their own friends and family. Users incorrectly concluded that they held unpopular views or were being given the cold shoulder. We are only at the beginning of understanding the implications of reasoning about algorithms. In the extreme case, it may be that whenever a software developer in Mountain View adjusts a parameter, someone somewhere suddenly starts to believe themselves unloved.

On learning that Facebook curates the News Feed, many users initially reacted with surprise, anger, or dissatisfaction. Notably, user satisfaction seemed lowest when users were informed that their friends were hidden. We found that people cared most about hidden *people*, while listings of hidden *content* had less emotional valence. This implies that when designing social networks, performing social filtering in addition to content filtering has real risks and triggers visceral reactions.

After learning about the algorithm with *FeedVis*, most of the same users who were initially upset became gradually more satisfied with the filtering process, implying that learning about the algorithm may be positive for opinions about the platform. When observing the algorithm outputs, participants

recognized that the algorithm hid many stories that were not directly related to them (e.g. others becoming friends). Although there was still hidden content our participants wished to see, when they reflected upon how the algorithm worked, most of them decided Facebook was “doing a pretty good job filtering out things.” (P7). We conclude that providing an understanding of why the algorithm exists and how the system’s results align with the user’s desires is a crucial next step when awareness of the algorithm is triggered.

We developed the *FeedVis* tool to reveal to users the contrast between algorithmically filtered and unfiltered results. This comparison between two algorithms (“filtered” vs. “most recent” or “show everything”) created algorithmic awareness and gives some understanding of how the algorithm works and where it might fall short. We believe that such tools have great potential. Tools like *FeedVis* could be extended to other domains or to demonstrate the performance of more than two algorithms. They could also be extended to allow users to create their own curation algorithms. Related “personally developed” algorithms have been explored in the past [13], and we argue that they will play an increasingly important role in the increasingly personalized online world.

What other insights might we draw from our findings to inform the design of technology? Designers of all types often struggle to determine what parts of a system’s operation should be made visible to their users. This study shows that the decision to promote a “secret sauce” or to highlight an otherwise hidden process is far more than marketing. Some designers prefer systems that operate as if by magic, delivering results without muddying the user experience with the details of a complicated process. In contrast, we suggest that enabling active engagement with the process shows users that an algorithm exists and gives them an important sense of agency — they are not controlled by an algorithm but are a part of one, and can have some influence on its results.

As we have suggested, our findings call on designers to consider algorithms an important factor when developing systems. As an example, consider the Twitter feed. At present, it is an unadulterated list of tweets. But if Twitter were to introduce an algorithmically curated feed, how should it make users aware of this change? In this scenario, we hope Twitter would consider the algorithm to be more than simply a way to manage information — but rather a way to offer users agency, control, and a deeper relationship with the platform itself. Algorithmic awareness is not a topic considered by most designers, but we believe it should be.

REFERENCES

1. Backstrom, L. News Feed FYI: A Window Into News Feed. <http://on.fb.me/1kz3Yak>. August 6, 2013.
2. Barocas, S., Hood, S., and Ziewitz, M. Governing algorithms: A provocation piece. In *Governing Algorithms: A Conference on Computation, Automation, and Control*. (2013).
3. Bennett, S. Social Media 2013: User Demographics For Facebook, Twitter, Pinterest And Instagram . <http://bit.ly/1j6g12u> March 19, 2013.

4. Bernstein, M. S., Bakshy, E., Burke, M., and Karrer, B. Quantifying the invisible audience in social networks. In *Proc. CHI 2013*, ACM Press (2013), 21–30.
5. Bilton, N. Facebook News Feed Draws More Criticism. *The New York Times*. March 5, 2013.
6. boyd, d. *It's Complicated: The Social Lives of Networked Teens*. New Haven: Yale University Press, 2014.
7. Brembs, B., Button, K., and Munaf, M. Deep impact: Unintended consequences of journal rank. *Frontiers in Human Neuroscience* 7, 291 (2013).
8. Diakopoulos, N. Sex, violence, and autocomplete algorithms, August 2013. <http://slate.me/1hMomGN>.
9. Diakopoulos, N. Algorithmic Accountability Reporting: On the Investigation of Black Boxes. *Tow Center for Digital Journalism* (2014).
10. DigitalShadow. <http://digitalshadow.com/>.
11. Floodwatch. <http://floodwatch.o-c-r.org/>.
12. Gaver, W. W., Boucher, A., Pennington, S., and Walker, B. Cultural probes and the value of uncertainty. *Interactions* 11, 5 (2004), 53–56.
13. Gilbert, E. Computing Tie Strength. *UIUC PhD dissertation* (2010).
14. Gilbreth, F. B., and Kent, R. T. *Motion study, a method for increasing the efficiency of the workman*. New York, D. Van Nostrand Company, 1911.
15. Gillespie, T. Can an Algorithm Be Wrong? <http://limn.it/can-an-algorithm-be-wrong/>.
16. Granka, L. A. The Politics of Search: A Decade Retrospective. *The Information Society* 26, 5 (2010), 364–374.
17. Hamilton, K., Karahalios, K., Sandvig, C., and Eslami, M. A path to understanding the effects of algorithm awareness. In *Proc. CHI EA 2014*, ACM Press (2014), 631–642.
18. Inc., F. How does News Feed decide which stories to show? <http://on.fb.me/1uuuZBL> Retrieved August 7, 2014.
19. Inc., F. What is News Feed? <http://on.fb.me/1q1111z> Retrieved September 15, 2014.
20. Introna, L., and Nissenbaum, H. Shaping the Web: Why the Politics of Search Engines Matters. *The Information Society* 16, 3 (2000), 1–17.
21. Kempton, W. Two Theories of Home Heat Control. *Cognitive Science* 10, 1 (1986), 75–90.
22. Kempton, W. Two Theories of Home Heat Control. *Cognitive Science* 10, 1 (1986), 75–90.
23. Kramer, Adam, G. J. E. H. J. T. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111, 24 (2013).
24. Litt, E. Knock, knock. Who's there? The imagined audience. *Journal of Broadcasting and Electronic Media* 56, 3 (2012), 330–345.
25. Lynch, K. *The Image of the City*. MIT Press: Cambridge, 1960.
26. Mandaro, L., and Guynn, J. Facebook talked ice buckets twice as much as Ferguson. *USA Today*. September 2, 2014.
27. McGee, M. EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close To 100K Weight Factors. <http://goo.gl/X4yOja>. August 16, 2013.
28. ProPublica Message Machine. <http://projects.propublica.org/emails/>.
29. Morris, M. R. Social networking site use by mothers of young children. In *Proc. CSCW 2014*, ACM Press (2014), 1272–1282.
30. Norman, D. *The Psychology of Everyday Things*. Basic Books: New York, 1988.
31. Nvivo Tool. <http://www.qsrinternational.com>.
32. Pariser, E. *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. Penguin Books, 2011.
33. Pasquale, F. Restoring Transparency to Automated Authority. *Journal on Telecommunications and High Technology Law* 9, 235 (2011).
34. Pegoraro, R. Facebook News Feed filtering can make friends vanish. *The Washington Post*. February 14, 2011.
35. Sandvig, C., Hamilton, K., Karahalios, K., and Langbort, C. Auditing algorithms: Research methods for detecting discrimination on internet platforms. In *Data and Discrimination: Converting Critical Concerns into Productive Inquiry* (2014).
36. Sanghvi, R. Facebook Gets a Facelift. <http://on.fb.me/1gd7z1I>. September 5, 2006.
37. Smith, C. 7 Statistics About Facebook Users That Reveal Why It's Such A Powerful Marketing Platform. <http://read.bi/194eLre>. November 16, 2013.
38. Steiner, C. Automate This: How Algorithms Came to Rule Our World. *Portfolio: New York* (2012).
39. Discrimination in online ad delivery. *Communications of the Association of Computing Machinery* (January 2013).
40. Making Facebook Less Infantile. *The New York Times*. August 8, 2012.
41. Valentino-Devries, J., Singer-Vine, J., and Soltani, A. Websites vary prices, deals based on users' information. *Wall Street Journal* (December 2012).