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

Motahhare Eslami, Aimee Rickman†, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong
Karrie Karahalios, Kevin Hamilton, Christian Sandvig‡
University of Illinois at Urbana-Champaign, †California State University, Fresno, ‡University of Michigan
{eslamim2, kvaccaro, aleyase2, avuong3, kkarahal, kham}@illinois.edu
†arickman@csufresno.edu, ‡csandvig@umich.edu

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 curation algorithm's existence at all. Initial reactions for these previously unaware participants were surprise and anger. We developed a system, *FeedVis*, to reveal the difference between the algorithmically curated and an unadulterated News Feed to users, and used it to study how users perceive this difference. Participants were most upset when close friends and family were not shown in their feeds. We also found participants often attributed missing stories to their friends' decisions to exclude them rather than to Facebook News Feed algorithm. 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's presence, however, algorithmic awareness led to more active engagement with Facebook and bolstered overall feelings of control on the site.

Author Keywords
Algorithms; Algorithm Awareness; News Feeds; Hidden Processes

INTRODUCTION
Today, algorithms curate everyday online content by prioritizing, classifying, associating, and filtering information. 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. An example of a prominent news feed today is Facebook News Feed, which in September of 2014, was viewed by on average 864 million daily active users [12]. This list of updating stories that appears front and center on Facebook home pages displays an algorithmically curated or filtered list of stories selected from a pool of all stories created by one's network of friends.

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, some teenagers added product names to their posts in an attempt to manipulate the algorithm and increase their posts' visibility [41]. Other users tended to block new mothers in their feed based on a false assumption that such women posted too many baby pictures, when in fact the prevalence of such images was determined by their popularity among users [27].

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. Algorithmic interfaces in Internet applications rarely include a clear enough feedback mechanism for users to understand the effects of their own actions on the system. Without such feedback, it can be difficult to assess the influence of either algorithm knowledge or ignorance.

To begin to address these issues, 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 [28, 29]. 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 understand why so few participants knew of the algorithm's existence, we conducted interviews to investigate Facebook use.

To assist us in these interviews, 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 difference between 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 and provided them with

CHI 2015, April 18 - 23 2015, Seoul, Republic of Korea

ACM 978-1-4503-3145-6/15/04$15.00
http://dx.doi.org/10.1145/2702123.2702556

an opportunity to modify the algorithm outputs to curate their desired News Feed. We discovered that strong initial negative reactions to the mere presence of an algorithmic filter 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 patterns 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 perceive and react to them. Cognitive science and human factors researchers study the mental models people create when they interact with machines and technology [30]. Designers develop new ideas by enacting probes that reveal interactions with hidden and uncertain aspects of people's lives [14]. Related efforts exist in architecture and urban planning, where architects create new spaces based on the study of how people perceive and navigate landscapes [24]. Finally, time and motion studies observe people conducting a task and extract any hidden patterns to find the most productive way to complete it [16]. 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 [23], including quantifying how perceived audiences compare to actual audiences and measuring otherwise invisible patterns of attention on social media [4].

Algorithms

Research attention has recently turned to algorithms as invisible and influential aspects of daily digital life. Many researchers have considered curation algorithms and argued that their effects are important while their operation is opaque [2, 19, 36]. For example, search algorithms structure the online information available to a society, and may function as a gatekeeper [18, 20]. And scientific journal rankings have been found to produce unwarranted perceptions about the importance of some articles over others [6].

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 [32, 39]. 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" [42]. Algorithms that select personalized advertisements have been found to unevenly distribute arrest record ads by race [40]. Controversy over Twitter Trends and accusations of algorithmic censorship of the tag #occupywallstreet throughout the Occupy Wall Street protests led to questions of whether a sorting algorithm can be wrong or unethical under some conditions [17, 35]. Some researchers have even studied unexpected results in the filtering of autocompletion text, finding some algorithms explicitly attempt to make moral judgements, such as removing terms deemed to be related to 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 to affect their online life [13, 26]. Designers at Ubisoft, a video game company, recently offered personalized inferences that can be made from Facebook profiles as a promotional device for a surveillance-themed game [10].

The prevalence of algorithmically generated feeds in social media such as Facebook News Feed has triggered discussions about the appropriateness of the curation algorithms employed. Some commentators are primarily concerned about friends that "vanish" from the platform [33], and others see an opportunity for profit linked to the position of posts [5]. While other work has attempted to reverse-engineer these algorithmic processes [19] or develop new summaries of algorithmic results [10, 13, 26], to our knowledge no researchers have developed systems to reveal to users the contrast between algorithmically manipulated and unfiltered results.

Case Study: Facebook News Feed

Launched in 2006 [37], the Facebook News Feed curation algorithm has attracted significant attention in recent years, particularly after a recent, controversial study of emotional contagion [22]. Facebook currently uses close to 100,000 factors to algorithmically choose the best stories from the large pool of potential stories for News Feed [25]. Although Facebook has stated it would change how it communicates updates to 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 the News Feed curation algorithm and what factors are associated with this awareness?

RQ2. How do users evaluate the curation of their News Feed when shown the algorithm outputs? Given the opportunity to alter the outputs, how do users' preferred outputs compare to the algorithm's?

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

STUDY DESIGN

To address the proposed research questions, we conducted a mixed-methods study consisting of three phases. First, participants visited our lab and completed a questionnaire and interview to measure algorithm awareness. At this time, we also collected participants' network size, News Feed stories and friends' stories to populate an interface for the next phase. In the second phase, participants used an application (FeedVis) to visualize the algorithm outputs, and we used a long form open-ended interview to discuss them. Third, we e-mailed participants two-to-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

At the beginning of the study, participants answered a demographic questionnaire including measures of their social media use. With one exception, all participants used Facebook at least once a day. 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 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 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 or a result of a filtering process. During this pre-assessment, we asked participants to use their Facebook accounts to log into our Facebook application, FeedVis. FeedVis extracted and collected the participant's network size, News Feed and their friends' stories. This collected information was used to generate a series of alternate views for the feed.

Main Interview: Algorithm Outputs Disclosure

After understanding the participants' existing News Feed knowledge, we presented them with a series of FeedVis feed views. Paging through these views revealed some algorithm outputs to the participants. If they were not already aware of the algorithm's existence, these views provided the first revelation of News Feed's algorithmic curation. These views were used as prompts in the interview so that participants could react to and discuss stories that actually appeared on their News Feed. As extracting all stories from an entire friend network is process-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 briefly describe the four FeedVis views.

The FeedVis Content View: Revealing Content Filtering

The Facebook algorithm shows a user a selection of stories chosen from the universe of all stories contributed by the people and pages that the user follows. In the first view, we aimed to show the user this universe of potential content, highlighting content that the algorithm excluded from 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*," included only the stories displayed on the user's News Feed. These stories were shown with a blue background. The left column, called "*All Stories*," showed every story posted from all the user's friends. 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. The content for the "Shown Stories" view was generated by querying user_id/home/user via the Facebook Graph API. It is important to note that "Shown Stories," while displayed on the user's News Feed, might not have been seen if the user did not scroll far enough. The content for the "All Stories" view is the union of friend_id/feed/ queries for each friend; we extracted all stories that the user would see if she went to a friend's page while logged in. We then used post_ids to determine whether those posts had appeared in the user's News Feed. To verify our operationalization of "Shown Stories," we asked participants if they remembered seeing randomly selected stories in this column. With a few exceptions, they did remember them.

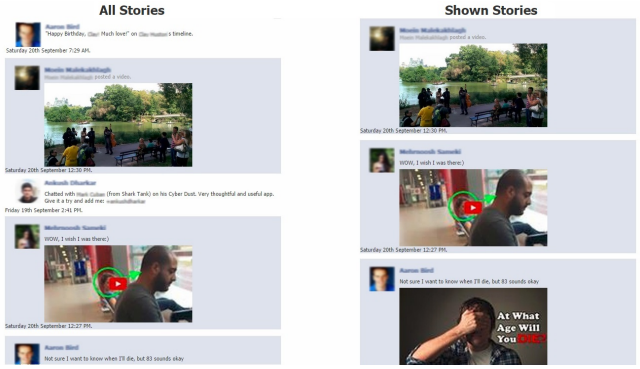

Figure 1. The Content View. Shown stories (in blue) occur across both columns, while the hidden stories (white) appear only in the left column as 'holes' in News Feed. Stories appear in reverse chronological order.

The FeedVis Friend View: Revealing Social Patterns

By filtering content, the Facebook algorithm also creates user perceptions about how other people use Facebook. We built a visualization, the Friend View, to help the user understand which users usually appear and which are hidden. This view divided the user's friends into three categories based on the proportion of each friend's stories that had appeared in the user's News Feed during the previous week: "rarely shown," "sometimes shown," and "mostly shown" friends (Figure 2).

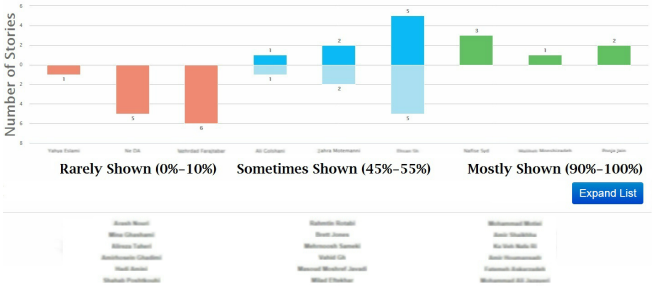

Figure 2. The Friend View. "Rarely shown" includes friends whose stories were mostly hidden (0%-10%) from the user. "Sometimes shown" includes friends who had roughly half of their posts (45%-55%) shown to the user. "Mostly shown" includes those friends whose stories were almost never filtered out (90%-100%) for the user. The number of the shown stories is displayed above the x-axis and the number of hidden stories is below the x-axis. The expand button augments the three category lists below the chart.

The FeedVis Friend & Content Rearrangement Views: Envisioning a Different Algorithm

After exploring the algorithm outputs, we wanted to gauge participants' desire to change them. We created two new

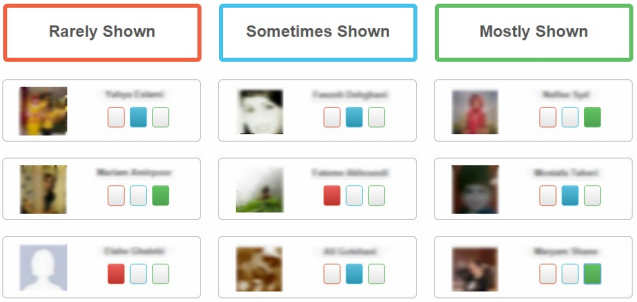

Figure 3. The Friend Rearrangement View. User can move friends between the categories by changing the color of a friend to the destination category's color.

views that invited participants to "tweak" their algorithm. The first view allowed adjustment based on story authorship, the second based on story content. First, the 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. Second, the Content Rearrangement View (Figure 4) randomly selected ten shown stories and ten hidden stories, then invited users to indicate whether they would have preferred a "shown" story to be "hidden" or vice versa. The lab portion of this study, including the pre-assessment, lasted one to three hours per participant.

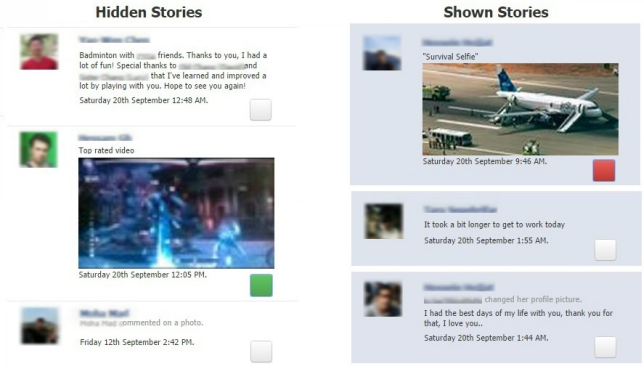

Figure 4. The Content Rearrangement View. User can move a story from its original category to the other by clicking the button beside each story.

Post-Assessment: Evaluating Algorithm Outputs Revelation

To understand the long-term consequences of revealing hidden aspects of a curation algorithm, we contacted participants via e-mail two to six months after conducting the study. We asked two questions and invited any additional comments participants wished to share. The questions were: (1) Has participation in our study resulted in more, less or no change in your satisfaction with Facebook News Feed? (2) Have you changed anything about how you use Facebook in light of what you learned in our study? (e.g., "I 'like' more posts now" or "I view posts using the 'Most Recent' setting instead of the 'Top Stories' setting.").

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 the Champaign, Illinois and surrounding area. 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 the University of Illinois and 19 people with other occupations such as homemakers, delivery persons, servers, 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 [3, 38].

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 scrolling 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 when so many others did not. Three 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 via one or two of the following common paths: inductively comparing feeds or deductively considering network size.

Inductively Comparing Feeds: Most Aware participants (n=12) compared the quantity of stories from different friends in their News Feed and felt they were seeing some friends' stories much more than 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). 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, 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 the algorithm was a basic, even obvious, element necessary to curate News Feeds and 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's existence? Were there any differences in Facebook usage associated with being aware or unaware of the 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. Some participants engaged with the algorithm outputs *passively* by, for instance, scrolling News Feed and reading the stories as they appeared. Some engaged with the algorithm outputs *actively*, 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 combined our interview material with data we extracted from each participant's Facebook account. We identified three passive and four active engagement features. Each feature was either mentioned by 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 in a relatively passive way — 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 prioritization in News Feed results in a greater proportion of filtered potential stories by the algorithm when the network is large[1].

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 may be more prone to active engagement with the algorithm outputs. They possibly explore more spaces on Facebook (such as options and settings screens) and may compare different aspects of the site. *Activity level*, a categorization of users as "listeners" (mostly reading the feed without posting a story), "light posters" (posting stories occasionally), or "heavy posters" (posting stories frequently), based on the participants' descriptions of their Facebook usage during the study. 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 which affect algorithm 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 importance, following a friend, hiding a story and making lists are some examples of these settings. Using any of these makes a user more actively engaged with the algorithm outputs because they are intentionally trying to change them. *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 possible existence of a filtering process).

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

We found a significant difference between the Aware and Unaware groups for all of the active engagement features by both

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

thematic and statistical analysis (Table 1). In terms of usage frequency, we found that all participants who used Facebook more than 20 times per day were aware of the curation algorithm. There was a significant difference in usage frequency between the Aware (M=27.18, SD=33.8) and Unaware participants (M=6.92, SD=5.79). Frequent users were on 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 allowed these participants to explore more stories, features and views (such as others' profile pages) than infrequent users. This exploration led to inductive feed comparisons and consequently new knowledge about News Feed 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

Participants unaware of the algorithm's existence included both posters and non-posters ("listeners"), but in contrast all 15 Aware participants were light or heavy posters. In Aware participants' discussions of their Facebook usage, we found the number of likes and comments on their own stories suggested the possibility of the existence of a filtering process. They found that their popular stories were shown in their friends' News Feeds more often: *"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] enough; and I experienced in my own post[s] [....] I think it probably has to do with the way Facebook presents [stories]"* (P23).

All six participants who did not apply any settings to adjust their News Feed content were unaware of the algorithmic curation 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 believed they could apply settings to adjust their News Feed if they were *"willing to invest the kind of time to find out how"* (P3), but did not invest this time.

There were seven participants involved in Facebook page/group management and all were aware of News Feed curation. 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 story 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 [7, 21], 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 knew 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 their reactions to an unfiltered alternative. Then we directed them to the Friend and Content Rearrangement Views, allowing them 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 completely new idea 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 interpersonally close enough. They were surprised to learn via FeedVis that that those hidden stories were likely removed by Facebook: *"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. For example, P37 asked:*"Do they choose what they think is the best for me to see? Based on what?"* This curiosity led them to wonder 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 was reflected on Facebook. This expectation was particularly likely for posts by family members; many participants stated that family members should appear in the "mostly shown" category in the Friend View: *"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).

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 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).

To better understand how participants' expected outputs compared to the actual algorithm outputs, we 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.

Despite the frustration in some initial reactions, more than half of the participants (n=21) came to appreciate 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[2]. 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). Although many participants were initially shocked, concerned or dissatisfied with the existence of a filtering algorithm, they concluded there were few 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). This suggests that while filtering is both needed and appreciated, a lack of awareness of the existence of the process leads to dissatisfaction.

From Algorithm Awareness to Future Behavior (RQ3)

During our initial discussions with Aware participants, we found their perceptions of the algorithm already affected their Facebook usage. They stated that awareness of the algorithm led them to actively manipulate their News Feed, using 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, [I'm] manipulating the system"* (P20). Others 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, a few Aware participants (n=4) tried to manipulate 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. For example, one participant suggested *"if you post a picture, without a comment, it's less likely to show up on your friends' News Feed"* (P21).

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. Of the 40 original participants, 30 responded[3].

Usage

Most of the follow-up participants (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, Aware and Unaware participants reported similar changes. The Aware participants specifically noted that FeedVis provided new information to them not available in the existing Facebook interface.

Manipulating the Manipulation: 21 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

[2]As 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.

[3]We 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.

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) because they preferred a time-sorted, unfiltered feed to Facebook's "Top Stories." A few stated that they *"tend to switch up between the 'Most Recent' setting and the 'Top Stories' setting"* (P14) to see both the trending and the chronological feed.

Ten 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 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 mentioned.

A few participants tried to make their own stories appear on more of their friends' News Feeds. For example, starting to like their own posts *"to give them more visibility"* (P28). Others modified their settings to limit who saw 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 [was] 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 the "Most Recent" story sorting setting, leaving them frustrated with the need to search for and understand settings. In an extreme case, one participant stopped using Facebook as she believed it was not straightforward with its users about News Feed curation: *"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 from the nine participants who reported no change in their usage, six 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 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 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 into the algorithm via FeedVis resulted in people feeling more knowledgeable about Facebook's algorithm and their satisfaction level with Facebook generally remained high.

LIMITATIONS

While our results are suggestive, we hope future research will employ a quantitative experimental design featuring a control group to better establish causal relationships between algorithmic awareness, its antecedents and consequences. Our study employed a non-probability sample and did not vary geographic diversity to match the US population, a dimension that may be important. And although the study was longitudinal, all behavioral assessment was based on self-reports.

As this study focused only on one instance of a curation algorithm, we do not know how far to generalize our conclusions. We suspect different dynamics exist in other contexts and even for other curation algorithms within Facebook (such as the algorithm that selects advertising).

Some FeedVis design decisions were influenced by the query limits in the Facebook API. At the time of our study, the Facebook API permitted 600 queries per minute. Therefore, it took longer to collect data for participants with larger friend networks. The size of a participant's network determined the time duration of the presented data for their FeedVis views.

While validating "Shown Stories" and "All Stories" feeds, we noticed that in a few cases, an expected story was not returned by the Facebook API. This finding has been reported by various developers and Facebook [11]. So this would not affect user perceptions, we used post_ids to ensure that "Shown Stories" were a subset of "All Stories" in the FeedVis views.

DISCUSSION

Users clearly benefit from awareness of an algorithmic curation process and likely from knowledge about how it works. Although algorithm awareness on Facebook was prompted by what we termed "active engagement," most users were not so engaged, and thus were not aware. We suspect that users are not aware of most curation, even when the presence of a filter appears obvious to those with a background in computing.

On Facebook, ignorance of the algorithm had serious consequences. Our participants used News Feed to make inferences about their relationships, wrongly attributing the composition of their feeds to the habits or intent of their friends and family. Users incorrectly concluded that friends had dropped them due to political disagreements or their unappealing behavior. In the extreme case, it may be that whenever a software developer in Menlo Park adjusts a parameter, someone somewhere wrongly starts to believe themselves to be unloved.

This conclusion draws our attention to more than just the danger of misunderstanding a filter. Users felt betrayed when discovering an algorithm that they were unaware of. Yet over time, knowledge about the algorithm increased satisfaction with the product. What are the best ways of alerting users to the presence of these processes? How much information about them is adequate to satisfy both the needs of effective interaction design and of principled, ethical use?

On the first question, as our study suggests that prolonged or passive use of Facebook did not correlate to knowledge of the algorithm at work, some form of direct intervention is required. An alert to the presence of an algorithmic process could take place external to the platform, as did ours, providing a sort of temporary x-ray or ombudsman's perspective into the composition of a normally seamless experience. This approach is consistent with that of interaction designers looking to create trust in critical systems such as voting machines; there, initial training and later auditing ensures trustworthy use [34]. Such an approach would also lend itself to continued "seamless" interaction with algorithmic media, avoiding the regular introductions of "seams" through explanations of algorithms that make interactions less fluid. However, while reverse engineering and explaining algorithms is promising, algorithms often use so many features that educating users about them is unlikely to be meaningful. And training through system help pages or blogs often falls short, both because users are unaware those resources exist and the resources provide too little information to be truly educational.

A different approach to alerting users to the presence and function of these algorithms could be integrated into routine use. Persistent, predictable feedback that enables users to understand a process has long been a staple of interaction design, and perhaps the introduction of new, more predictable capabilities for "tweaking" one's feed, or one's appearance in other feeds, achieves awareness of algorithms without sacrificing fluidity and dependability. Providing a visual narrative for algorithmic processes has the potential to educate users without revealing technical specifications or intellectual property. We argue that providing this kind of feedback requires trusting the user, but we believe all are capable of better understanding how their digital environments work.

In either case, work remains to be done on just how much information is enough to satisfy the needs of trustworthy interaction, civic good, and, pragmatically, the protection of proprietary interest. Our study provides some starting points for this work. It shows that users respond differently to revelations about different features of the algorithm (e.g., friend-based vs. story-based filtering). 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; related "personally developed" algorithms have been explored in the past [15], 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 often struggle to determine what parts of a system's operation should be made visible to 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 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 that they are not controlled by an algorithm but are a part of one, and can have some influence on its results. Indeed, the algorithm can offer users agency, control, and a deeper relationship with the platform itself.

In conclusion, given the recent rise of concerns over the ethical and social consequences of opaque algorithms in search, news and other applications, it is high time for interaction designers to bring their own approaches to the conversation. Arguments for algorithm transparency by ethicists or journalists may strike more pragmatic developers as unrealistic, and product satisfaction or popularity as a primary standard for evaluating success will strike others as inadequate for ensuring sound civic roles for these powerful platforms. Future work in this area will require contributions from many angles.

Like many complex infrastructures, our algorithmic platforms reflect the influence of economic interests, empirical and design research, and competing foundational assumptions about collective living. If the best systems have achieved success through careful integration of such disparate approaches into the design process, certainly our algorithms deserve the same.

ACKNOWLEDGMENTS

This work was funded in part by the Office for the Vice Chancellor of Research Interdisciplinary Innovation Initiative at the University of Illinois. We would like to thank Roshanak Zilouchian, John Lee, Mary Pietrowicz and Amin Sadeghi for early feedback on this work.

REFERENCES

1. Backstrom, L. News Feed FYI: A Window Into News Feed. http://on.fb.me/1kz3Yak.
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.
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*.
6. Brembs, B., Button, K., and Munaf, M. Deep impact: Unintended consequences of journal rank. *Frontiers in Human Neuroscience 7*, 291 (2013).
7. Collins, A., and Centner, D. How people construct mental models1. *Cultural models in language and thought* (1987), 243.
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. Facebook Developers. https://developers.facebook.com/bugs/576745515760299/.
12. Facebook newsroom. Company Info. http://newsroom.fb.com/company-info/.
13. Floodwatch. https://floodwatch.o-c-r.org/.
14. Gaver, W. W., Boucher, A., Pennington, S., and Walker, B. Cultural probes and the value of uncertainty. *Interactions 11*, 5 (2004), 53–56.
15. Gilbert, E. Computing Tie Strength. *UIUC PhD dissertation* (2010).
16. 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.
17. Gillespie, T. Can an Algorithm Be Wrong? http://limn.it/can-an-algorithm-be-wrong/.
18. Granka, L. A. The Politics of Search: A Decade Retrospective. *The Information Society 26*, 5 (2010), 364–374.
19. 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.
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. 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).
23. Litt, E. Knock, knock. Who's there? The imagined audience. *Journal of Broadcasting and Electronic Media 56*, 3 (2012), 330–345.
24. Lynch, K. *The Image of the City*. MIT Press: Cambridge, 1960.
25. McGee, M. EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close To 100K Weight Factors. http://goo.gl/X4yOja.
26. ProPublica Message Machine. http://projects.propublica.org/emails/.
27. Morris, M. R. Social networking site use by mothers of young children. In *Proc. CSCW 2014*, ACM Press (2014), 1272–1282.
28. Facebook Inc. How does News Feed decide which stories to show? https://www.facebook.com/help/166738576721085.
29. Facebook Inc. What is News Feed? https://www.facebook.com/help/210346402339221.
30. Norman, D. *The Psychology of Everyday Things*. Basic Books: New York, 1988.
31. Nvivo Tool. http://www.qsrinternational.com.
32. Pasquale, F. Restoring Transparency to Automated Authority. *Journal on Telecommunications and High Technology Law 9*, 235 (2011).
33. Pegoraro, R. Facebook News Feed filtering can make friends vanish. *The Washington Post*.
34. Pieters, W. Explanation and trust: what to tell the user in security and ai? *Ethics and information technology 13*, 1 (2011), 53–64.
35. Sandvig, C., Hamilton, K., Karahalios, K., and Langbort, C. Re-centering algorithms. In *Governing Algorithms: A Conference on Computation, Automation, and Control.* (2013).
36. 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).
37. Sanghvi, R. Facebook Gets a Facelift. http://on.fb.me/1gd7z1I.
38. Smith, C. 7 Statistics About Facebook Users That Reveal Why It's Such A Powerful Marketing Platform. http://read.bi/194eLre.
39. Steiner, C. Automate This: How Algorithms Came to Rule Our World. *Portfolio: New York* (2012).
40. Discrimination in online ad delivery. *Communications of the Association of Computing Machinery* (January 2013).
41. danah boyd: The kids are all right. http://www.pbs.org/wgbh/pages/frontline/media/generation-like/danah-boyd-the-kids-are-all-right/.
42. Valentino-Devries, J., Singer-Vine, J., and Soltani, A. Websites vary prices, deals based on users' information. *Wall Street Journal* (December 2012).