The Illusion of Control: Placebo Effects of Control Settings

Kristen Vaccaro[1], **Dylan Huang**[1], **Motahhare Eslami**[1], **Christian Sandvig**[2], **Kevin Hamilton**[1], and **Karrie Karahalios**[1,3]
[1]University of Illinois, [2]University of Michigan, [3]Adobe Research
{kvaccaro, dphuang2, eslami2, kham, kkarahal}@illinois.edu, csandvig@umich.edu

Figure 1. Users engage in complex sensemaking around control settings, without discovering that some of the controls display random tweets.

ABSTRACT
Algorithmic prioritization is a growing focus for social media users. Control settings are one way for users to adjust the prioritization of their news feeds, but they prioritize feed content in a way that can be difficult to judge objectively. In this work, we study how users engage with difficult-to-validate controls. Via two paired studies using an experimental system – one interview and one online study – we found that control settings functioned as placebos. Viewers felt more satisfied with their feed when controls were present, whether they worked or not. We also examine how people engage in sensemaking around control settings, finding that users often take responsibility for violated expectations – for both real and randomly functioning controls. Finally, we studied how users controlled their social media feeds in the wild. The use of existing social media controls had little impact on user's satisfaction with the feed; instead, users often turned to improvised solutions, like scrolling quickly, to see what they want.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

ISBN 978-1-4503-5620-6/18/04. . . 15.00
DOI: https://doi.org/10.1145/3173574.3173590

Author Keywords
Social media, sensemaking, control settings, placebo effect

INTRODUCTION
Much past work has explored the design of control panels and control settings, particularly in hardware and safety-critical domains (e.g., control panels ranging from washing machines and ovens [14, 42] to aircraft and nuclear reactors [41, 46]). However, surprisingly little has been said about the design and interaction with control settings in graphical user interfaces (with a notable exception of [33]), beyond traditional user interface design practice and through encouraging sensible general design choices as analogues to such hardware designs [42]. This gap in knowledge is surprising and problematic as control settings and control panels are omnipresent in people's everyday use of technology.

Social media platforms have added control settings to increase user satisfaction and feelings of control (particularly around privacy), but also as a general feedback mechanism to increase user engagement [43]. While personalization of news and social media feeds was proposed early (e.g., Fishwrap and Excite [8, 36]), such news feeds often featured highly manual personalization, for example, choosing to "remove politics." Much of this personalization is now done automatically. While users can provide feedback and control, such changes are often less direct (e.g., "see less like this"), may change only some elements of the feed (rather than more visible changes like removing an entire section) and there may be delays before users see the impact of their choices. In this work, we consider the question of whether and how users validate control settings.

Control settings on social media differ from those on traditional applications, because they can be very challenging to judge objectively. A user might know precisely what to expect if they click "*add subtitles*" on a video service, select a "*large*" font size on their web browser, or filter a list of products to "*less than $50*." But social media applications can be much more difficult to interpret. If, for example, a user clicks "*Hide post: See fewer posts like this*," (as on Facebook), it is not obvious what "posts like this" means or how clicking it will change the user's news feed. When the effectiveness and functionality of social media controls are difficult for users to validate, the question arises: do these controls increase satisfaction because of their operation or because of their mere presence?

In this paper, we present the results of two experiments exploring this topic. Both use systems that show users their Twitter news feeds with control settings we developed. An interview study offered real and random algorithmic controls to explore user sensemaking (Figure 1). An online study included an additional no-control-setting condition to test for a placebo effect. Through these studies, we uncover a placebo effect for control settings, where users are more satisfied with a social media news feed when a control setting is present, whether it functions with a real algorithm or simply displays a random subset of the news feed.

MOTIVATION AND RELATED WORK

One of the most validated and replicated phenomena is that of placebo effects – that the mere presence of a "cure" (even if it is an inert, non-operational pill) heals people [38]. Medicine [27], psychology [56], and even marketing [54] have well established the strength and power of this effect. While placebo effects are more often controlled for than studied themselves, a placebo can cause "objective (structurally and functionally measurable) changes in physiological functioning" [24]. While many different theories for the mechanism of the placebo effect exist [24], an important factor is that the intervention should not be trivial for the patient to assess. Thus, while many control settings (like a price filter) may be straightforward to evaluate, given the difficulty of objectively judging the functionality of many social media controls, we raise the question of whether they too may function as placebos.

Such a placebo effect test has rarely been investigated in computing. Researchers have explored the potential of *benevolent deception* [1], either for entertainment [37] or to help users learn [53]. However, as far as we know only Springer et al. have directly studied a placebo effect, where users interacted with a supposed "emotion meter" and believed that a completely random system was working well [55]. This motivates our first research question, where we explore whether there is a placebo effect for control settings, particularly for difficult-to-judge algorithmic control settings typical of social media. In the remainder of the paper, we frequently refer to these control settings simply as controls, and a control panel as a grouping of such controls.

RQ1: Is there a placebo effect for control settings — that is, are users more satisfied with a social media feed when a control is present, whether the control works or not?

Understanding how users make sense of automated systems is important not only to design for accessibility for a range of users during routine operation, but to help users navigate malfunction or unexpected outcomes. Indeed, as Lucy Suchman demonstrated in her classic study at Xerox, design for user experience of system breakdowns — when even the most trusted human-machine relationship weathers difficulty — is often crucial to the creation of humane and lasting systems [57]. As a result of the important role sensemaking plays in user experience and design, many researchers have explored sensemaking around automated systems, in domains from thermostats and home sensors [58, 63] to self tracking devices [30, 47, 64] and other personal data [62] to classification algorithms [31].

Recent work has even begun to explore how users make sense of randomly automated systems that are deceptively presented as functional [55]. This work (along with others that do not directly test non-functional operations) has shown that users are so likely to trust an algorithmic system that when making sense of the system, they will disregard evidence of failures by taking personal responsibility for them, or assuming an algorithmic omniscience that extends beyond the system [18, 55, 62]. In this work, we similarly study how sensemaking processes differ when some control settings function randomly (but are presented as real), in the domain of social media.

Many researchers have explored sensemaking around social media and news feeds [7, 12, 18, 20, 48]. French and Hancock identified the metaphors people hold for news feeds, from rational assistants to unwanted observers [20]. Raider and Gray found many users believed social media systems engage in complex reasoning to infer their preferences and show them relevant posts [48]. Eslami et al. found that, driven by their theories of how news feeds are algorithmically controlled, users engage in complex behaviors like *liking* and then immediately *hiding* posts to try to achieve the feed they want [18].

Researchers have also highlighted the significance of experiencing violated expectations on social media. Bucher's interviews with Facebook users revealed that this can lead to a feeling of creepiness [7]. DeVito et al.'s study of folk theories of algorithmic curation of the Twitter news feed showed that users' resistance to algorithmically curated news feeds is often driven by violated expectations [12]. In our second research question, we focus on expectation violation with algorithmic controls of social media – additionally studying how (if at all) these responses differ when a control functions randomly:

RQ2: How do people make sense of controls, particularly when the control functionality violates their expectations?

Finally, despite the prevalence of control settings on social media, studies of how people are able to find them and how likely they are to use them have primarily occurred within the narrow domain of privacy controls [4, 26, 28, 61, 65]. Numerous researchers have noted that user takeup of privacy controls is low [2, 3, 23]. Some have sought to identify why users have difficulty finding and understanding privacy controls [61]. We seek to explore the use of controls other than privacy control settings. How do participants use Twitter's existing controls? Does control usage (specifically, usage of the follow and unfollow controls central to Twitter's news feed) impact user satisfaction? And when imagining their ideal news feed, how do users envision using controls (if at all) to create it?

RQ3: How do users currently edit – and envision editing – their news feed? Does more frequent control use make users more satisfied with their current feed?

In order to address these research questions, we conducted two experiments, designed to complement each other – one an in-person interview, one online. Both used similar systems, based on a common design that paired the user's own Twitter news feed with control settings they could use to influence the algorithmic composition of their feed. Figure 2 illustrates the two system designs used for these experiments. In some cases, the controls worked "correctly," updating the news feed to capture one implementation of the concept for the control setting. In others the controls worked "randomly," instead showing a consistent but random subset of the users' news feed. In the online experiment, a third, baseline condition was added, where no control setting was present. Users answered questions about their experience with the controls as well as the resulting feed to help us address these research questions.

SYSTEM DESIGN

We conducted two related experiments allowing the user to control their own social media news feed. For each, the user interacted with the control settings (if available) to try to improve their news feed and answered questions about their experience. Popular social network Twitter (330 million monthly active users [60]) was selected because it allowed us to abide by Terms of Service while accessing users' unfiltered news feeds.

Figure 2(a) illustrates the core features common to both experimental systems: the user login, the news feed and any available controls. Users log in and are shown their own news feed on the left. The number of controls on the right depends on the experiment; the online experiment had a single or no control, the interview experiment had five control settings. When users interact with the controls, their news feed immediately updates. Instant updates enable Schön's reflection-in-action – users who immediately see the impacts of their choices can better engage in sensemaking around the outcomes [50].

Control Settings for Social Media

Users encounter many types of control settings on social media platforms. While social media controls can be complex and granular, most systems allow a user to control 1) who sees them and 2) who and what they see. The first overlaps with privacy controls, where a user can decide whether they can be discovered, tagged in photos, and whether their content is public. Many control settings also allow the user to control what they see on their own news feed: controls of ads (both whether ads are personalized and the interests and location history those are based on), controls of friends (e.g., follow or friend), more granular controls of what content from friends is seen (e.g., block, mute), and finally controls that both provide feedback but also potentially impact the feed (e.g., like). For this experiment, we designed controls that would be difficult to judge objectively (e.g., popularity) but would revise the feed immediately when a participant modified the control setting.

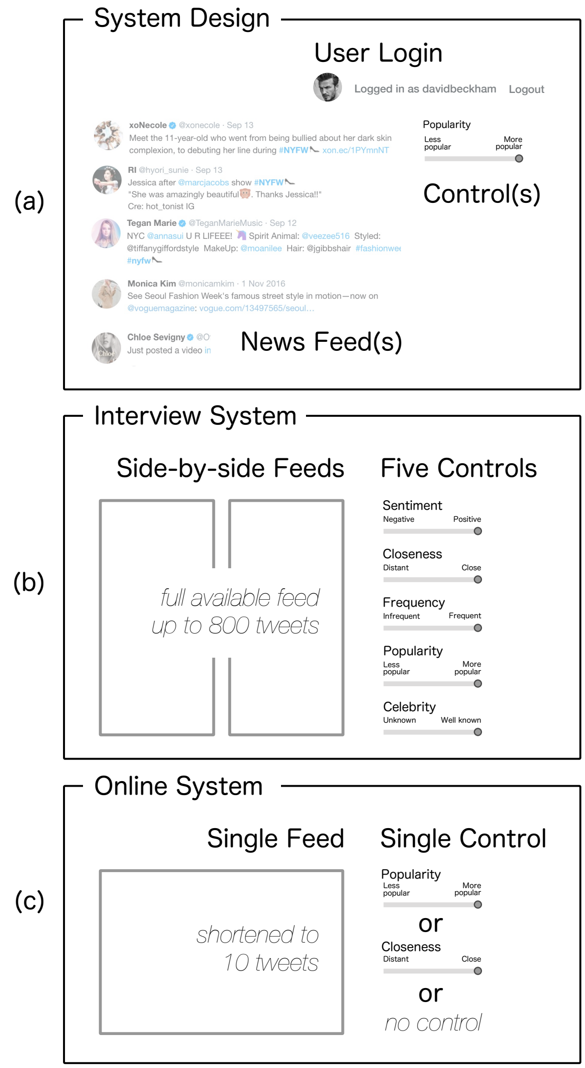

Figure 2. Two experiments used similar system designs. The overview (a) shows the core features: a user login, the user's own news feed, and controls. The interview system design (b) featured five controls and side-by-side feeds, while the online system (c) had a single control and feed.

While few social networks provide controls for the kind of summary features (like popularity) we developed, they often provide access to the results on alternate views like Instagram's "Explore" tab. External tools also provide similar controls [66].

Control settings must also be useful for the particular social media platform that they control. Twitter functions as a social network, but also as a news media site [32]. To capture both of these aspects, we developed control settings for accounts but also for tweet content. Controls were roughly split between tweet content (sentiment, popularity) and accounts (poster frequency, closeness and celebrity). With these controls, we tried to maximize four factors: 1) *Ecological validity*: these controls are typical of those found in social media, 2) *Value*: pilot tests suggested users care about these concepts, 3) *Non-triviality*: these controls cannot be trivially falsified, unlike some alternatives (e.g., a "media" control that toggles the presence of photos where failures are trivial to detect) and 4) *Data support*:

there is data from the platform to support implementation of a working control. While controls for personalized content like "Things you might like" are appealing, the Twitter API does not provide data (e.g., a user's like history) to support it. Controls were iterated on repeatedly to meet these goals.

Real and Random Control Setting Algorithms

We briefly outline how each control's values were calculated:

Popularity used a simple metric: the number of times the post was retweeted.

Sentiment was calculated using LIWC [45]. Using a positive and negative vocabulary, tweets were scored by the total of number of positive minus negative words.

Frequency measured how often an account posted, using the average number of times they posted per month, over the length of the account's existence.

Celebrity measured how famous an account was using its number of followers. The scores increased for users with more than 1000 followers, as the average user (who has posted in the last month) has about 700 [35]. The score used a log scale: $C = max(log(x) - 3, 0)$ for x followers.

Closeness approximated tie strength to capture how personally close a user was to the account they followed. This measure was based on features used for Twitter in the WeMeddle system [22]. Using the current API, we were able to compute a score based on the number of followers, number and sentiment of direct messages, most recent direct message, and how recently the account was followed, where the closeness score increased inversely for number of followers and recency of following.

The algorithms we developed for these control settings may not be optimal. For example, sentiment might be improved by using other dictionaries [25] or including more complex features like word length [6]. However, they captured the concepts well enough, as shown through numerous pilot tests conducted to ensure that the algorithms produced useful results and in the validation phase of the interview experiment.

Since one of the purposes of the experiments was to understand whether it matters if controls work or not, each control had both a "real" and a "random" algorithm for computing scores. For each user, we collected the 800 most recent tweets and all accounts they follow, using the Twitter API. For each tweet or account, we computed values for our controls using these data. Once all non-random values were computed, "random" values were assigned by drawing randomly from this distribution. In this way, non-uniform distributions (e.g., of popularity) could be captured for the random controls.

Interface Details

Each control setting was presented as a slider (Figure 1b), with a title (e.g., celebrity) and each end labeled (e.g., unknown and well-known). We provided neither text nor verbal explanations of the function of each control setting. This more closely emulated existing social media systems, where it is left to the user to guess what impact a *like*, *block* or even *follow* will have on their overall news feed and social media experience. Each slider used a coarse 5-point granularity, where the ticks formed equally sized bins of accounts or tweets (depending on control type) for the logged in user. Thus the scores for an account would be relative only to the other accounts the user followed, so @MerriamWebster (~500k followers) might be the most well-known celebrity for one user, but a less known celebrity for another following more famous accounts.

In the face-to-face experiment the user viewed two news feeds side-by-side (Figure 2b). One news feed was controlled by the real scores for each control, while the other was controlled by the random scores. The left-right ordering of which feed was controlled by real or random controls was randomized automatically. Seeing both sides update simultaneously helped users compare the changes and engage in reflection-in-action [50].

In the online experiment, on the other hand, the user viewed a single feed, shortened to 10 tweets (Figure 2c). As the user interacted with the single control (when a control was present), the news feed would update either using either the real *or* the random scores for that control type (e.g., popularity). The ordering of when users encountered real, random and no control setting for a given control type were similarly randomized.

EXPERIMENTAL DESIGN

We investigated our research questions using a mixed methods approach. We designed two complementary studies, the first a qualitative and quantitative laboratory study conducted face-to-face (the "interview experiment"), and the second a quantitative online experiment. The qualitative portion of the lab study allowed for our sensemaking investigation. In the online study, we scaled our study to a larger number of more geographically diverse participants, using only quantitative survey measures to understand participants' preferences. The online study also included an additional no-control-setting condition as a baseline measure to study the placebo effect.

Both studies investigated the effects of introducing control settings to a news feed on self-reported user experience. In both studies we experimentally manipulated the independent variable of control settings by varying their presence and function in three possible levels: real controls present, random controls present, and in the online experiment only: no control settings present. We then measured the dependent variables user satisfaction, feelings of control, and in the in-person lab study only: sense-making approach. Dependent variables were measured after each treatment: both studies were posttest-only designs. Different control settings (using the algorithms described above) were employed, but in order to focus on our research questions were not treated as a factor; both experiments are therefore one-way and test the presence and function of control settings, conceptualized as one factor. Experiments are within-subjects, as participants interacted with every condition. Treatment order was counterbalanced by randomization.

In the interview experiment, we addressed RQ3 by first capturing detailed data via surveys and interview questions on participants' control setting use and their satisfaction with their existing feeds, prior to any experimental intervention. We then tested two treatments: real controls and randomly functioning controls. Semi-structured interviews allowed par-

Frequency of Use		% of Users Read	Post		Usage Yesterday	
Multiple times	/day	84%	16%		Once	9%
Once		6	9		2-3	31
Few times	/week	6	16		4-5	28
Once		—	6		6-10	19
Few times		3	19		>10	9
Once	/month	—	9			
Less than once		—	19			
Never		—	6			

(a) (b)

Figure 3. Reported Twitter Use. Participants report their overall use (a), reading Twitter frequently, though many post less frequently (if ever). Most reported using Twitter several times yesterday (b).

ticipants to share their sensemaking process for each type in an open-ended way and to address RQ2 – how they responded to violated expectations. Finally, the interview experiment included a manipulation check to confirm that the experimental conditions were distinguishable: we told participants that some controls functioned randomly, and participants were asked to distinguish the real settings from the random ones.

The online experiment supplemented this, specifically testing whether there is a placebo effect for controls, to address RQ1. Since a placebo effect is measured in comparison to a no-intervention baseline (i.e., an inert pill making people healthier than no treatment at all), the online study included an additional baseline "no-control-setting" treatment. The effects of these interventions were measured with Likert scale survey questions that participants answered for each condition. The online experiment also featured a larger, more geographically representative sample than the interview study.

The two studies were designed to complement each other via a triangulation design using the convergence model [10]. Qualitative results from the in-person interviews were merged with the quantitative findings of the lab and online experiments in the interpretation stage. While the control panels differ in the number of controls (and number of conditions), both studies provide independent complementary evidence about sensemaking and the placebo effect. In both experiments, the user interacted with controls to "make their favorite version of their news feed for leisurely browsing (e.g., while drinking coffee on a weekend morning)" and answered questions about their experience. We describe each in more detail below.

INTERVIEW EXPERIMENT

The interview experiment was designed to study sensemaking practices (RQ2), capture current social media and control use (RQ3), and confirm the real controls worked well enough for users to distinguish them from the randomly functioning ones.

Participants

Although the interview experiment employed a non-probability convenience sample, participants were recruited with varied appeals designed to increase demographic diversity. We advertised the study with flyers in public locations (libraries, coffee shops, golf courses, etc.) and with online advertisements (including newsletters, news boards, etc.). The results include interviews with 32 participants (44% male, 6% chose not to answer) from the Champaign-Urbana area. Participants were diverse (19% Black, 19% Asian, 6% Hispanic/Latino, 3% Multiracial and 53% White), though the sample skewed young (59% under 25).

The selection criteria for participation was reading Twitter at least once per day; most participants reported reading Twitter at least once per day (Figure 3), though they reported posting to Twitter less frequently. Participants also reported the number of times they used Twitter yesterday and on a typical day in free text. Distributions were quite similar and were collapsed from real values to the groupings reported in Figure 3b.

Participants were paid $10/hour for the duration of their interview. The full interview typically ran an hour and a half (M=1 hour 38 min, SD=20 min).

Study Tasks

Part I: Pre-Experiment Survey and Feed Exploration

The first segment of the interview experiment focused on understanding users' current use of social media. After the user logged in, they completed a survey on their current social media usage and control usage on social media. Users were then shown 10 tweets randomly selected from their home timeline and asked to explain why they would or would not want to read each. Next, the main system displayed the most recent tweets from their home timeline (Figure 2b) and users pointed out any tweets they especially liked or disliked.

These tasks served two functions: first, they qualitatively captured how actively each user curates their news feed already. Second, these activities served to surface the participants' preferences and help focus their attention on what they like/dislike and on how they might use the controls in the next stage.

Part II: Control Sensemaking

The second stage of the experiment focused on understanding how users make sense of controls. Participants saw two news feeds side-by-side (Figure 2b), one controlled by the real scores for the controls, the other controlled by the random scores. Comparing left and right as they updated simultaneously helped users think more deeply about the controls.

Participants first performed the tasks for this segment focusing on the left-hand feed, then repeated the same process focusing on the right-hand feed. The left-right ordering of real and random sides was randomized by the system.

Participants were asked to use any controls they liked to create their favorite version of their feed for leisurely browsing. The interviewers conducted a semi-structured interview, asking questions as the participants engaged with the controls. When the participant was satisfied with their news feed, they read it for a few minutes, before completing a survey about their experience and satisfaction with the feed they created. Finally, they were asked to explicitly compare the results on the two sides and identify which of the feeds they preferred (left or right) and summarize their thoughts about the controls.

Part III: Validation and Debriefing

Participants might be as satisfied with random controls as with real ones if the "real" controls were designed *badly* enough that users found their outcomes indistinguishable from random changes. To address this potential issue, we included a final manipulation check segment in the interview study.

After users answered questions about their ideal feed, they were debriefed with questions about what they thought the purpose of the study was. They were then told the true purpose of the study. Users were shown a duplicate interface, where the left hand column always updated at random. Users were told that one of the news feeds was updating at random as they chose different controls and were asked to identify which feed was updating at random, again using any controls they liked.

Measures

The questions in this study were presented 1) in a pre-experiment survey, 2) in an after-task feed survey, and 3) as open-ended questions within the semi-structured interview. The pre-experiment survey (given at the beginning of the experiment in Part I) included demographic questions and questions about participants' typical social media use. The after-task feed survey was given twice, once after the user had experienced each of the two feeds (real and random) in Part II. In the after-task feed survey, Likert scales assessed the dependent variables, descriptives, and manipulation checks for the quantitative experiment in the lab study. These included:

Satisfaction A scale of three questions assessed satisfaction with the news feed: *I am satisfied with the final news feed I saw for leisurely browsing*, *I enjoyed browsing this news feed*, and *I would like to use an interface like this in my day-to-day browsing of Twitter*.

Feelings of control One question was used as a manipulation check to ensure that users felt more control in the experimental conditions when a control setting was present: *I feel in control of my news feed*. We assess control with a single item to avoid priming participants to focus on our manipulation.

While the questions were developed ad hoc for this experiment, they were iterated on to be easily interpreted, unambiguous and to capture a variety of aspects of satisfaction (with the feed, interface, and imagined future use).

Open-ended questions within the semi-structured interview spanned the length of the experiment, capturing dependent variables, descriptives, and manipulation checks such as:

Sensemaking approach When the users engaged with the controls, the open-ended questions included: *What do you expect with that option?*, *How do you think it's doing?*, *What are you paying attention to, to tell whether it's working?*, and *Is there anything you're surprised to see? Or that you think is missing?* to capture sensemaking approach and satisfaction.

Current control use In addition to questions in the survey, the debriefing included questions about what the user's ideal or "dream" feed would include and *If you wanted to try to achieve this with Twitter's controls now, how would you do so?*

Identification of random feed Users were asked in the debriefing at the end of the study if they noticed anything in particular about the options they used or the way the feeds changed. In the validation phase, these included asking directly *Which do you think is updating randomly?* and *How did you work that out?* to capture whether (and how) the user identified the "random" feed.

ONLINE EXPERIMENT

The online experiment supplemented the interview experiment to specifically determine whether there is a placebo effect for controls (RQ1). We added a no-control-setting condition (not present in the interview study), to allow comparisons between real and random controls and a no-control-setting baseline.

Participants

The online experiment recruited participants from Amazon Mechanical Turk to supplement the interview experiment with more participants and participants from more diverse locations. Although it also employs a non-probability convenience sample, the online experiment includes surveys from 106 participants from 30 states, covering every region of the continental US. Participants (50% male, 1% chose not to answer) ranged between 20 and 78 years old ($M = 35$, $SD = 10$).

The full online experiment was designed to last less than 30 minutes. Participants in this experiment were paid $4, for an approximate rate of $8/hour, though online participants were not individually timed.

Study Tasks

For the online experiment, the system was significantly simplified; participants interacted with a single control or the baseline no-control-setting condition. They viewed a single feed at a time (Figure 2c), rather than two side-by-side feeds as in the interview. For each such feed, participants answered a short survey about their experience with the control and feed.

After logging in, users were shown a brief demographic survey. They were then shown their news feed (abbreviated to the most recent 10 tweets) and a single control. From the full five controls developed for the interview study, the online study focused on only two of the controls: popularity and closeness. These two were selected to capture the measures focused on tweet content (popularity) and those focused on accounts (closeness).

Each page participants viewed constituted one experimental condition. Each page showed a news feed and a control setting (closeness or popularity), which was then manipulated (real/random/no-control-setting). Participants viewed pages in sets of three focusing on a given control type (Figure 4), where each of the three pages showed one of the three experimental conditions (real/random/no-control-setting) in a randomized order. Participants saw each control and condition three times, to address any possible learning effects, for 18 total experiences. Although this could be analyzed as a 2 x 3 full (or partial) factorial design, for clarity we did not treat the control type (popularity vs. closeness) as a factor in this study.

As in the interview experiment, users were asked to create their favorite version of their feed using the controls (if available). The participant was then shown a short survey assessing satisfaction and self-efficacy for each page.

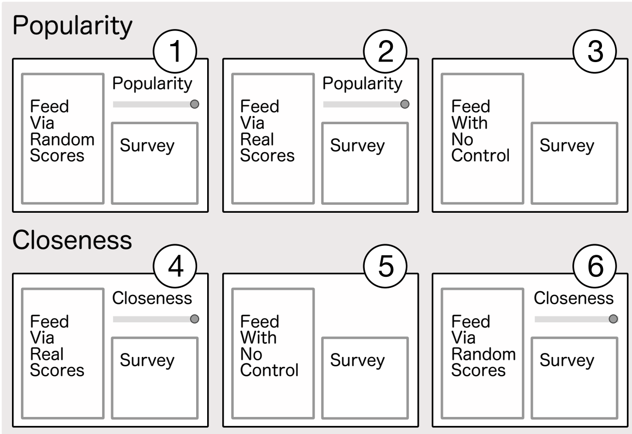

Figure 4. Online Experiment Design Example. Participants experienced each control type (popularity and closeness) and condition (real, random, no-control-setting), where ordering was randomized within each control type. Participants would experience a series like this three times.

Measures

The survey questions in this experiment took two forms– attention check and dependent measures. One question in each survey was an attention check, common in Mechanical Turk surveys. For example, an attention check might read "*Please select the option* strongly agree." Of the 18 attention check questions, participants needed to answer 15 correctly or their results were omitted from the analysis. The other four questions on each survey were the same Likert scales assessing satisfaction and feelings of control as were described in the measures section for the interview experiment.

ANALYSIS

Interviews were analyzed by the first author using iterative open coding. Survey questions were analyzed statistically in the following way: First, the consistency of user responses was assessed with Cronbach's alpha to test for possible learning effects. Second, the reliability of the scale for satisfaction was assessed using item-total correlation scores. The manipulation check and scale results were then analyzed using the Kruskall-Wallis test to account for the ordinal values of the Likert scale results. Finally, if the Kruskall Wallis test indicated an effect, a post hoc Dunn test was used to identify which conditions (real, random or no-control-setting) differed.

RESULTS

We identify a placebo effect for control settings, where users are more satisfied with their news feed when a control setting is offered, whether it works or not. Users experienced violated expectations when engaging with both real and random controls, often taking responsibility for system failures. Importantly, we also show that more frequent manipulation of control settings on Twitter did not increase users' satisfaction with their news feed. Instead, users often refered to improvised means of customizing the experience.

RQ1: Placebo Effect of Control Settings

We found users were more satisfied with feeds when a control was present, regardless of whether it functioned randomly. In addition, the interview study confirmed that the "real" controls worked well enough for users to distinguish them from ones with a random algorithm, once made aware of this distinction.

Online study results

To address possible learning effects with our experimental system, participants were shown each control and condition pair three times. Users answered consistently, with a mean Cronbach alpha of 0.79 ($\alpha = 0.71 - 0.89$), indicating that users scored a given control type/condition pairing (e.g. popularity/real) similarly each time they encountered it.

The Kruskall Wallis ANOVA test revealed no effect ($H = 0.85, p = 0.36$) across the popularity and closeness controls for the question "*I feel in control of my news feed*." So popularity and closeness responses were collapsed for a Kruskal-Wallis test comparing the real, random, and no-control-setting conditions for the same question. A strong effect was found ($H = 94.47, p \ll 0.00001$), so Dunn post hoc tests were computed for each pair of conditions (Table 1). The scores functioned as a manipulation check, and indicated that users did in fact feel greater control when a control panel was present.

Item-total correlation scores for the three satisfaction questions confirmed the reliability of the scale (popularity $r = 0.66 - 0.78$, closeness $r = 0.73 - 0.86$). The satisfaction scale had significant differences between the popularity and closeness controls ($H = 3.92, p = 0.047$), where user satisfaction was higher for popularity controls. So popularity and closeness results are reported separately. Computing Kruskal-Wallis scores for the three conditions revealed a significant difference for both popularity ($H = 33.85, p \ll 0.0001$) and closeness ($H = 13.85, p < 0.001$). Post-hoc Dunn's tests uncovered an effect for real/none and random/none and no significant difference for real/random pairs (Table 1). That is, users were more satisfied and felt more in control when using a system with either real or random controls than no control setting, and furthermore are no less satisfied with random controls.

Interview study results

Users were similarly satisfied with both real and random controls in the interview study, but users could identify the randomly-controlled feed once made aware of the distinction.

Participants completed the same scale in the after-task feed survey after interacting with the real and randomly controlled feeds. Again, we found no significant difference between user satisfaction with the real and random controls ($H = 1.31, p = 0.25$), though this study lacked a no-control-setting baseline.

	Real & Rand	Real & None	Rand & None
In Control	$Z = 0.89$ $p = 0.38$	$Z = 8.83$ $p \ll .00001$	$Z = 7.94$ $p \ll .00001$
Satisfaction Popularity	$Z = 0.74$ $p = 0.46$	$Z = 5.37$ $p \ll .00001$	$Z = 4.63$ $p \ll .00001$
Satisfaction Closeness	$Z = -0.80$ $p = 0.43$	$Z = 2.75$ $p < .01$	$Z = 3.55$ $p < .001$

Table 1. Dunn *post hoc* test results. Users are more satisfied and feel more control when using either real or random controls than no control, but there is no significant difference between real and random controls.

In addition, when we asked users if they noticed anything about the controls in our debriefing, no participant identified that the controls on one column were working randomly. One participant came close, noting, "*The right side seemed more like it actually was doing what you picked. The left side was kinda random. Maybe not random, but it wasn't as what I wanted as the right*" (P29). It was more common for participants to note that one side seemed to work *better* than the other ($n = 6$), but some preferred the randomly-generated news feed that they saw as a less literal interpretation of the concepts:

> *It seemed like the right side did exactly what I wanted it to do when I picked a certain filter and the left side didn't necessarily do exactly what I wanted it to do but did a better job of kind of inferring what I wanted, in a way* (P10).

The majority of participants mentioned some other unrelated feature of the controls, e.g., controls they particularly enjoyed, controls they wished we would add, or sharing concerns about the concept of filtering social media in general ($n = 22$).

One concern was that participants might not notice a difference between real and random controls, if the real controls were designed badly. To address this, the validation task asked users to identify which feed was being updated randomly, after being made aware that one feed would do so. Four participants incorrectly identified which feed updated randomly; 28 identified the randomly updating news feed correctly.

RQ2: Addressing Violated Expectations

Designing algorithms (particularly machine learning algorithms) to be interpretable by users is an active and challenging topic of research [15, 34]. Even the simplest algorithms can be difficult to interpret from their output, particularly if users do not have clear expectations of what the algorithm will do. Even with our simple implementations (e.g., popularity measured only using retweets), users developed sophisticated theories of what the controls do. While it is outside the scope of this paper to detail the rich folk theories users developed for the controls, we highlight how users addressed violated expectations.

Past work has shown that even for non-functioning automated systems, users will disregard opposing evidence to believe they are working, even generating excuses for the system, like "*I did not give it enough to work with*" [55]. We find similar results, where users developed explanations for unexpected results. Interestingly, this was true for both real and random controls – users also had their expectations violated by the functional control settings, and excused them in similar ways. We report the ways users addressed violated expectations in the interview stage, reporting the number of users who expressed a similar statement for the real and randomly functioning controls.

In fact, one method of addressing violated expectations was substantially more common with the real controls: stating the controls were fighting each other ($n_{real} = 10$, $n_{rand} = 1$). As one user described this:

> *I mean, I'm trying to imagine what the sliders do and thinking of how I've maybe created a conflicting environment where I want things that are more popular but less frequent. I don't know, something that's popular because a lot of people retweeted or favorited it or liked it, or whatever they call it now. Maybe that doesn't always match with things that are infrequent, I guess* (P25).

The user takes responsibility, in this case, for having created a demand that is impossible to satisfy.

Most ways of addressing violated expectations were used for both real and random controls, though often slightly more frequently for random controls. Users often updated their own theory about how the control worked ($n_{real} = 15$, $n_{rand} = 16$) or explained what might be happening ($n_{real} = 5$, $n_{rand} = 11$):

> *It may just be a function of the tweets it has to sort through that only 'famous' people were tweeting in my feed. Of the people that I follow, only the famous ones were tweeting an hour ago or whatever* (P20).

A few explanations were quite common. Users explained tweets by conditioning their description, like describing someone as relatively unknown for a celebrity ($n_{real} = 10$, $n_{rand} = 16$). When seeing a comedian in the feed after moving the sentiment control setting to negative, one participant noted, "*So this, for example, it's a comedian but, this comedian pretends he's like Kim Jong Un from North Korea. So, it's pretty dark humor, a lot of negative things*" (P9). Users also noted that different people might interpret a result differently ($n_{real} = 3$, $n_{rand} = 8$), for example when looking at the results for more well-known celebrities: "*Probably, if you're a sports fan — I'm not really a sports fan [...] but then if someone saw this picture, they would have known who this is, but I don't*" (P23). Some even second guessed their own evaluation of accounts, content or even their own preferences ($n_{real} = 7$, $n_{rand} = 10$): "*I feel like they're not really less popular? Maybe they are and I just, I noticed them more than they actually post... Like since I like them I think they post a lot but they really don't*" (P2).

Finally, a number of users took responsibility when the system violated their expectations, either saying it was their fault for not being able to understand the controls ("*Are we close? I don't know, I'm horrible at this one!*" (P5)) ($n_{real} = 2$, $n_{rand} = 7$), or taking responsibility for doing something outside the system, "*Oh, they're still there. I should probably just unfollow them if I don't like them so much*" (P15) ($n_{real} = 4$, $n_{rand} = 7$).

While our results have shown that users are more satisfied when a control setting is present, we have also discovered that users will take responsibility – for asking too much of the system, for not being able to understand, for not doing something they should have already – when surprised or dissatisfied with the results, whether the control works or not. We next study how this plays out on Twitter now, capturing how participants use some of the existing controls on Twitter and how that impacts their satisfaction with their feed.

RQ3: Current and Envisioned Use of Twitter's Controls

We find a wide range in how often people use following controls on Twitter, but more editing does not make people enjoy their feed more. When users address how to be more satisfied with their news feed, they imagine using existing controls, but also a number of improvised approaches.

Control Use	% Users who control who they follow by			
	Add Users	Avg % of Tweets Liked	Remove Users	Avg % of Tweets Liked
Daily	3%	30%	6%	35%
Weekly	16	54	—	—
Monthly	69	50	31	57
Yearly	9	56	38	52
>1/Year	3	80	13	45
Never	—	—	13	53

Figure 5. Current control use and impact on user satisfaction. Most add users at least a few times per month, though they remove less frequently. Updating who they follow has little impact on news feed satisfaction.

Impact of Use of Existing Controls on News Feed Satisfaction

Interview results indicated that users typically do use existing controls frequently, but that higher frequency use of controls has little impact on users' satisfaction with their existing feed. Results draw from the interview study's pre-experiment survey on current social media usage. Twitter has an extensive control panel available to users, including fine grained controls that range from muting whole accounts to muting particular words from just notifications from people you don't follow for 24 hours. Our survey focused on use of follow and unfollow controls, which form the core basis of the Twitter news feed.

Figure 5 shows participants' current use of Twitter's existing controls for adding and removing followers from the news feed. There is a wide variety of current control use, ranging from daily use to several users who report never removing users once they follow them. Users typically update who they follow by adding followers at least a few times per month.

To understand how users' control use impacts their satisfaction with their news feed, we use data from the pre-experiment feed exploration. The fraction of 10 randomly selected tweets that users reported wanting to see is used as a measure of their satisfaction with their existing news feed. We find that more frequent control use (curating who the user follows) has no impact on the fraction of tweets the user wants to see, either by adding ($r = -0.22, p = 0.22$) or removing ($r = -0.03, p = 0.86$) friends. This suggests that use of existing controls has little impact on users' satisfaction with their existing feed.

How Users Envision Controlling What They See

We also studied how users envision controlling what they see. When addressing how to be more satisfied with their news feed, users point to existing controls, but also describe many improvised behaviors like scrolling fast or simply choosing different social media platforms to control what they see. These results include data from 1) users' reflections (while using our system) about how they achieve their goals on Twitter now and 2) an interview question where users described their "dream feed" (i.e., their ideal version of their news feed) and then how they would create it using Twitter.

Existing controls. Many participants did envision using existing controls offered by Twitter to create their dream feed. The most frequently referred to control was following and unfollowing users ($n = 20$). Users typically thought of this as quite straightforward. For example to create their dream feed: "*Follow more accounts. And unfollow the accounts that tweets [sic] a lot. I think that's it*" (P19). Users also referred to existing controls including muting or blocking users ($n = 4$), lists ($n = 4$), or Twitter's "Show me the best Tweets first" ($n = 2$).

Improvised controls. However, many users also described using other Twitter affordances or behaviors to view what they wanted. Several described using the explore tabs (which include search, trending, and moments) to see current and popular tweets ($n = 8$), so that their news feed did not need to include those: "*You can press search and [Twitter] brings you to a whole 'nother page. It's like trending stuff on there, and that's usually, that's how I find out most of my news* " (P29).

Others describe going directly the profile pages of accounts they particularly like or care about to see the full feed ($n = 4$), whether a news account or a personal friend. For example, one user described visiting friends' profiles to "*read their tweets directly there, instead of in a feed with anybody else*" (P25).

A few rarely look at their news feed and rely primarily on direct messages from friends to see what they want ($n = 2$). As one participant noted, all his friends share tweets. And whether they are shared humor or sharing opinions about important events, "*It's like for me. It's geared towards me*" (P21).

Finally, a few users who did not envision many opportunities for themselves to create their dream feed, instead used their own behavior – scrolling quickly – to avoid tweets they didn't want to see ($n = 3$). As one described it,

> *Oh, how would I do that now? I think just by scrolling quickly when I see stuff. Or when I see... I said before, like 50 retweets in a row, I'm just like "Okay, scroll, scroll scroll, scroll, scroll"* (P7).

Finally, the largest group of participants described controlling what they see by the choice of social network itself ($n = 10$). They viewed selecting a social network as a form of controlling who and what they saw, by following different people on different platforms. Several noted that they used other platforms for closer friends (e.g., "*I don't have them on Snapchat 'cause I'm not a close enough friend*" (P3)), but Twitter for more distant friends or celebrities:

> *I don't go on Twitter to look at my friends, so that's Instagram or Snapchat. Twitter is very much to look at the opinions of people that I typically don't know but that have a bigger platform* (P30).

Some tied this to how social norms differ across platforms, "*I'm Facebook friends with people I hate and all that. But with Twitter, it's more like you have the choice to follow who you follow*" (P4). Others focused on the core difference in news feed prioritization between social networks ("*Sometimes that stuff gets buried [...] 'cause I feel like with Facebook I miss a lot*" (P11)), tying that to the importance of Twitter for news: "*I don't usually follow newspapers on Facebook or Instagram [... But] for breaking news and stuff I go to Twitter right away*" (P9).

So while users understood the available controls, they described use of a rich set of improvised solutions to be more satisfied with their feeds. This is somewhat surprising, given the extensive control settings that Twitter offers.

CONCLUSIONS, QUESTIONS, AND NEXT STEPS

Trust, Violated Expectations and Ethics of Controls

In this work, we uncovered a placebo effect for control settings on social media. While this placebo effect might suggest adding non-functional controls – or even just a quick and dirty implementation of the controls eventually desired – to increase user satisfaction, we argue here that we must look deeper.

Non-functional controls might suffice if social media were used solely for entertainment (as was likely the case for most of our participants). However, research has shown that social media has far greater impact, e.g., on organizing social movements [59] and news access [52]. Reliance on a platform such as Twitter in the context of a violent political conflict might bring high stakes to the questions of reliability and trustworthiness of control settings or of the algorithm itself.

The fact that, in our work, users encountered violated expectations for both the real and the random controls suggests a potential for breakdown of trust even with the real controls. Indeed, Kizilcec has shown that when users' expectations are violated, they trust systems less — unless provided some transparency [29]. While our study did not explicitly measure trust, it is an important factor when considering the impact of such a placebo effect. It is already well-established that users default to considering automated systems trustworthy [16] and may be liable to *over*trusting systems [39, 40, 49].

For controls in particular, the likelihood that users accept defaults has potential negative outcomes, including forming inaccurate representations of important relationships or world events that, when acted upon, lead to misunderstanding, strained relationships, social alienation, or worse [17, 51]. With social media playing a growing role in civic and political life, surely it will become ever more important to establish not only whether these platforms are delivering what they promise, as in the growing area of work on algorithmic accountability, but how they establish trustworthiness for users. Relatedly, explorations of the ethics of misrepresentative control settings might fit well within emerging conversations about the ethics of algorithmic systems [5, 11, 13, 44].

Social Context and Platform

No control setting can exist apart from the system it controls. To complicate our call for transparency in interface design, we reference Matthew Fuller who argues:

> *Whenever an interface promises to make something 'clear' or speaks of allowing something to work in just the simplest way possible, it must first of all be assaulted with questions, yawns and scripts rather than rewarded with immediate identification* [21].

In our case, for example, we consider the question: what are the control settings not being built? One could argue that a social media platform would *never* build the control settings people truly want (e.g., removing ads or promoted content). For example, one user noticed the lack of ads on our system, "*In my normal feed [...] honestly on every page there's two or three either sponsored, or it says ad. This took them all out, which is amazing*" (P31). But since this runs counter to the economic interests of the platform, control settings to remove advertisements seem unlikely.

As a platform, Twitter is popular, but it differs from other social networks: the default is public posts, following is unidirectional (allowing connections with celebrities) and Twitter has important size limitations. This may limit generalization of the results in some important ways. Perhaps use of existing controls would be more closely tied to satisfaction in less fast-paced platforms or ones that focus more on personal friends, whose presence past work has shown to be important [19].

This study focused on user satisfaction with news feeds without regard for the social and historical context of the user and platform. This raises potential for future study that accounts for such contexts more explicitly – or for the possibility that user satisfaction may be a secondary feature in some contexts.

Rethinking Control Panel Design

Our work showed the presence of a placebo effect but not its mechanism. Are people more satisfied when a control is present simply because they enjoy exerting control? Or is it the change to their daily feed, and the accompanying sense of surprise or serendipity of what they're seeing, that makes users more satisfied? One user described his news feed goal as: "*I pretty much only wanna see basketball, but then having these little things, of like the most basic thing to know about weather, fashion, business...*" (P28). The added spice of such atypical results may be why users enjoyed random controls.

Understanding the mechanisms that determine this effect will likely also impact what good control setting design practices should look like. However, our study suggests that traditional methods of study (like the cognitive walkthrough) are not sufficient. Particularly as much of the control users have is intertwined with other functionality with important social implications (e.g., the "like" button), it is also an open question how users weigh the desire for control against other important social media functions.

In addition, users' rich set of improvised approaches to seeing what they want raises questions. Given the dexterity users show in "*inhabit[ing] a stream of multilayered information*" on Twitter [9], perhaps it should be unsurprising that these approaches exist. Nevertheless, we wonder what it is about manipulation of the news feed through scrolling techniques, patterns of use, or other improvised solutions that seems to bring more satisfaction for users (particularly given the rich set of existing controls). Better understanding of how users employ improvised manipulation of news feeds, and why, could support designers of feed composition algorithms in creating products that create trust and support long-term use.

ACKNOWLEDGEMENTS

We would like to thank Michael Kruepke and Wai-Tat Fu for their helpful discussions of experimental design. This work was supported by NSF grant CHS-1564041.

REFERENCES

1. Eytan Adar, Desney S Tan, and Jaime Teevan. 2013. Benevolent deception in human computer interaction. In *Proc. CHI*. DOI: http://dx.doi.org/10.1145/2470654.2466246
2. Naveen Farag Awad and Mayuram S Krishnan. 2006. The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. *Management Information Systems Quarterly* (2006). DOI: http://dx.doi.org/10.2307/25148715
3. Susan B Barnes. 2006. A privacy paradox: Social networking in the United States. *First Monday* 11, 9 (2006). DOI:http://dx.doi.org/10.5210/fm.v11i9.1394
4. Andrew Besmer and Heather Richter Lipford. 2010. Moving Beyond Untagging: Photo Privacy in a Tagged World. In *Proc. CHI*. DOI: http://dx.doi.org/10.1145/1753326.1753560
5. Reuben Binns. 2017. Algorithmic Accountability and Public Reason. *Philosophy & Technology* (2017), 1–14. DOI:http://dx.doi.org/10.1007/s13347-017-0263-5
6. Samuel Brody and Nicholas Diakopoulos. 2011. Cooooooooooooooollllllllllll!!!!!!!!!!!!!!: using word lengthening to detect sentiment in microblogs. In *Proc. EMNLP*. Association for Computational Linguistics.
7. Taina Bucher. 2017. The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms. *Information, Communication & Society* (2017). DOI: http://dx.doi.org/10.1080/1369118X.2016.1154086
8. Pascal Chesnais, Matthew Mucklo, and Jonathan Sheena. 1995. The Fishwrap Personalized News System. In *Proc. 2nd International Workshop on Community Networking*. DOI:http://dx.doi.org/10.1109/CN.1995.509583
9. Kate Crawford. 2009. Following you: Disciplines of listening in social media. *Continuum* 23, 4 (2009), 525–535. DOI: http://dx.doi.org/10.1080/10304310903003270
10. John W Creswell and Vicki L Plano Clark. 2011. Designing and Conducting Mixed Methods Research. (2011).
11. John Danaher. 2016. The threat of algocracy: reality, resistance and accommodation. *Philosophy & Technology* 29, 3 (2016), 245–268. DOI: http://dx.doi.org/10.1007/s13347-015-0211-1
12. Michael A DeVito, Darren Gergle, and Jeremy Birnholtz. 2017. Algorithms Ruin Everything: #RIPTwitter, Folk Theories, and Resistance to Algorithmic Change in Social Media. In *Proc. CHI*. DOI: http://dx.doi.org/10.1145/3025453.3025659
13. Nicholas Diakopoulos and Michael Koliska. 2017. Algorithmic Transparency in the News Media. *Digital Journalism* (2017). DOI: http://dx.doi.org/10.1080/21670811.2016.1208053
14. Alan Dix, Janet Finlay, Gregory Abowd, and Russell Beale. 2004. *Human-computer interaction*. Pearson Education. DOI: http://dx.doi.org/10.1007/978-1-4899-7993-3_192-2
15. Finale Doshi-Velez and Been Kim. 2017. Towards a rigorous science of interpretable machine learning. https://arxiv.org/abs/1702.08608, (2017).
16. Mary T Dzindolet, Scott A Peterson, Regina A Pomranky, Linda G Pierce, and Hall P Beck. 2003. The Role of Trust in Automation Reliance. *International Journal of Human-Computer Studies* 58, 6 (2003), 697–718. DOI: http://dx.doi.org/10.1016/S1071-5819(03)00038-7
17. Motahhare Eslami, Amirhossein Aleyasen, Roshanak Zilouchian Moghaddam, and Karrie Karahalios. 2014. Friend Grouping Algorithms for Online Social Networks: Preference, Bias, and Implications. In *Proc. SocInfo*. Springer. DOI: http://dx.doi.org/10.1007/978-3-319-13734-6_3
18. Motahhare Eslami, Karrie Karahalios, Christian Sandvig, Kristen Vaccaro, Aimee Rickman, Kevin Hamilton, and Alex Kirlik. 2016. First I Like It, Then I Hide It: Folk Theories of Social Feeds. In *Proc. CHI*. DOI: http://dx.doi.org/10.1145/2858036.2858494
19. Motahhare Eslami, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. 2015. I always assumed that I wasn't really that close to [her]: Reasoning about Invisible Algorithms in News Feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 153–162. DOI:http://dx.doi.org/10.1145/2702123.2702556
20. Megan French and Jeff Hancock. 2017. What's the Folk Theory? Reasoning About Cyber-Social Systems. http://dx.doi.org/10.2139/ssrn.2910571, (2017). DOI: http://dx.doi.org/10.2139/ssrn.2910571
21. Matthew Fuller. 2003. *Behind the Blip: Essays on the Culture of Software*. Autonomedia. p. 107.
22. Eric Gilbert. 2012. Predicting Tie Strength in a New Medium. In *Proc. CSCW*. DOI: http://dx.doi.org/10.1145/2145204.2145360
23. Eszter Hargittai and Alice Marwick. 2016. "What Can I Really Do?" Explaining the Privacy Paradox with Online Apathy. *International Journal of Communication* 10, 0 (2016). http://ijoc.org/index.php/ijoc/article/view/4655
24. Anne Harrington. 1999. *The placebo effect: An interdisciplinary exploration*. Harvard University Press.
25. Clayton J Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proc. ICWSM*.
26. Sanjay Kairam, Mike Brzozowski, David Huffaker, and Ed Chi. 2012. Talking in Circles: Selective Sharing in Google+. In *Proc. CHI*. ACM, 1065–1074. DOI: http://dx.doi.org/10.1145/2207676.2208552

27. Ted J Kaptchuk, John M Kelley, Lisa A Conboy, Roger B Davis, Catherine E Kerr, Eric E Jacobson, Irving Kirsch, Rosa N Schyner, Bong Hyun Nam, Long T Nguyen, and others. 2008. Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. *BMJ* 336, 7651 (2008), 999–1003. DOI: http://dx.doi.org/10.1136/bmj.39524.439618.25

28. Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W Reeder. 2009. A nutrition label for privacy. In *Proc. SOUPS*. ACM, 4. DOI: http://dx.doi.org/10.1145/1572532.1572538

29. René F Kizilcec. 2016. How much information?: Effects of transparency on trust in an algorithmic interface. In *Proc. CHI*. DOI: http://dx.doi.org/10.1145/2858036.2858402

30. Simon Knight, Theresa Anderson, and Kelly Tall. 2017. Dear Learner: Participatory Visualisation of Learning Data for Sensemaking. In *Proc. International Learning Analytics & Knowledge Conference*. ACM, 532–533. DOI:http://dx.doi.org/10.1145/3027385.3029443

31. Todd Kulesza, Margaret Burnett, Weng-Keen Wong, and Simone Stumpf. 2015. Principles of Explanatory Debugging to Personalize Interactive Machine Learning. In *Proc. IUI*. ACM. DOI: http://dx.doi.org/10.1145/2678025.2701399

32. Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. 2010. What is Twitter, a Social Network or a News Media?. In *Proc. WWW*. DOI: http://dx.doi.org/10.1145/1772690.1772751

33. Clayton Lewis and John Rieman. 1993. *Task-centered user interface design*.

34. Zachary C Lipton. 2016. The mythos of model interpretability. *arXiv preprint arXiv:1606.03490* (2016).

35. Ryan MacCarthy. 2016. The Average Twitter User Now has 707 Followers. https://kickfactory.com/blog/average-twitter-followers-updated-2016/. (June 2016).

36. Paul Maglio and Rob Barrett. 2000. Intermediaries personalize information streams. *Commun. ACM* 43, 8 (2000). DOI:http://dx.doi.org/10.1145/345124.345158

37. Joe Marshall, Steve Benford, and Tony Pridmore. 2010. Deception and magic in collaborative interaction. In *Proc. CHI*. ACM. DOI: http://dx.doi.org/10.1145/1753326.1753410

38. Daniel E Moerman. 2002. *Meaning, Medicine, and the "Placebo Effect"*. Cambridge University Press. DOI: http://dx.doi.org/10.1017/CBO9780511810855

39. Bonnie M Muir. 1994. Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics* 37, 11 (1994), 1905–1922. DOI: http://dx.doi.org/10.1080/00140139408964957

40. Bonnie M Muir and Neville Moray. 1996. Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics* 39, 3 (1996), 429–460. DOI: http://dx.doi.org/10.1080/00140139608964474

41. Randall J Mumaw, Emilie M Roth, Kim J Vicente, and Catherine M Burns. 2000. There is more to monitoring a nuclear power plant than meets the eye. *Human factors* 42, 1 (2000), 36–55. DOI: http://dx.doi.org/10.1518/001872000779656651

42. Donald A Norman. 2002. The Design of Everyday Things. (2002). DOI: http://dx.doi.org/10.15358/9783800648108

43. Will Oremus. 2016. Who Controls Your Facebook Feed. http://www.slate.com/articles/technology/cover_story/2016/01/how_facebook_s_news_feed_algorithm_works.html. (January 2016).

44. Frank A Pasquale. 2011. Restoring transparency to automated authority. https://ssrn.com/abstract=1762766, *Journal on Telecommunications and High Technology Law* (February 2011).

45. James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. *The development and psychometric properties of LIWC2015*. Technical Report.

46. Ivan Poupyrev, Desney S Tan, Mark Billinghurst, Hirokazu Kato, Holger Regenbrecht, and Nobuji Tetsutani. 2001. Tiles: A Mixed Reality Authoring Interface. In *Proc. Interact*.

47. Aare Puussaar, Adrian K Clear, and Peter Wright. 2017. Enhancing Personal Informatics Through Social Sensemaking. In *Proc. CHI*. ACM. DOI: http://dx.doi.org/10.1145/3025453.3025804

48. Emilee Rader and Rebecca Gray. 2015. Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed. In *Proc. CHI*. DOI: http://dx.doi.org/10.1145/2702123.2702174

49. Paul Robinette, Wenchen Li, Robert Allen, Ayanna M Howard, and Alan R Wagner. 2016. Overtrust of robots in emergency evacuation scenarios. In *Human-Robot Interaction (HRI)*. IEEE. DOI: http://dx.doi.org/10.1109/HRI.2016.7451740

50. Donald A Schön. 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books. DOI: http://dx.doi.org/10.1080/07377366.1986.10401080

51. Rajiv C Shah and Christian Sandvig. 2008. Software defaults as de facto regulation the case of the wireless Internet. *Information, Community & Society* 11, 1 (2008), 25–46. DOI: http://dx.doi.org/10.1080/13691180701858836

52. Elisa Shearer and Jeffrey Gottfried. 2017. News Use Across Social Media Platforms 2017. http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/. (2017).

53. Jaeeun Shim and Ronald C Arkin. 2013. A taxonomy of robot deception and its benefits in HRI. In *Systems, Man, and Cybernetics (SMC)*. IEEE, 2328–2335. DOI: http://dx.doi.org/10.1109/SMC.2013.398
54. Baba Shiv, Ziv Carmon, and Dan Ariely. 2005. Placebo Effects of Marketing Actions: Consumer May Get What they Pay For. *Journal of Marketing Research* (November 2005). DOI: http://dx.doi.org/10.1509/jmkr.2005.42.4.383
55. Aaron Springer, Victoria Hollis, and Steve Whittaker. 2017. Dice in the Black Box: User Experiences with an Inscrutable Algorithm. *AAAI Spring Symposium Series* (2017).
56. Steve Stewart-Williams and John Podd. 2004. The Placebo Effect: Dissolving the Expectancy versus Conditioning Debate. *Psychological Bulletin* 130, 2 (2004), 324. DOI: http://dx.doi.org/10.1037/0033-2909.130.2.324
57. Lucy A Suchman. 1987. *Plans and situated actions: The problem of human-machine communication*. Cambridge university press.
58. Peter Tolmie, Andy Crabtree, Tom Rodden, James Colley, and Ewa Luger. 2016. "This has to be the cats": Personal Data Legibility in Networked Sensing Systems. In *Proc. CSCW*. ACM. DOI: http://dx.doi.org/10.1145/2818048.2819992
59. Zeynep Tufekci. 2017. *Twitter and tear gas: The power and fragility of networked protest*. Yale University Press.
60. Twitter. 2017. Q1 2017 Letter to Shareholders. http://files.shareholder.com/downloads/AMDA-2F526X/5234607751x0x939175/D7BAFE57-DCBD-42E9-9909-7F587047FCED/Q117_Shareholder_Letter.pdf. (April 2017).
61. Kami Vaniea, Lujo Bauer, Lorrie Faith Cranor, and Michael K Reiter. 2012. Out of sight, out of mind: Effects of displaying access-control information near the item it controls. In *Privacy, Security and Trust (PST)*. IEEE, 128–136. DOI: http://dx.doi.org/10.1109/PST.2012.6297929
62. Jeffrey Warshaw, Tara Matthews, Steve Whittaker, Chris Kau, Mateo Bengualid, and Barton A Smith. 2015. Can an Algorithm Know the Real You?: Understanding People's Reactions to Hyper-personal Analytics Systems. In *Proc. CHI*. ACM. DOI: http://dx.doi.org/10.1145/2702123.2702274
63. Rayoung Yang and Mark W Newman. 2013. Learning from a learning thermostat: lessons for intelligent systems for the home. In *Proc. Ubicomp*. ACM. DOI: http://dx.doi.org/10.1145/2493432.2493489
64. Rayoung Yang, Eunice Shin, Mark W Newman, and Mark S Ackerman. 2015. When fitness trackers don't 'fit': end-user difficulties in the assessment of personal tracking device accuracy. In *Proc. Ubicomp*. ACM. DOI: http://dx.doi.org/10.1145/2750858.2804269
65. Xuan Zhao, Niloufar Salehi, Sasha Naranjit, Sara Alwaalan, Stephen Voida, and Dan Cosley. 2013. The many faces of Facebook: Experiencing social media as performance, exhibition, and personal archive. In *Proc. CHI*. ACM. DOI: http://dx.doi.org/10.1145/2470654.2470656
66. Ethan Zuckerman. 2017. Who Filters Your News? Why we built gobo.social. https://medium.com/mit-media-lab/who-filters-your-news-why-we-built-gobo-social-bfa6748b5944. (2017).