

Interventions that Inform

Designing Communications that Highlight News Filter Bubbles
and Provide Strategies for Combatting their Negative Effects.

Thank you for the guidance in this thesis to my advisors Peter, for his invaluable knowledge in the field of machine learning and its junction with design, and Stacie, for her infinitely constructive feedback and encouragement, and for teaching me how to navigate ambiguity with perseverance and grit. A special thanks to my parents and my colleagues at the CMU School of Design for the endless supply of inspiration and support throughout graduate school.

Interventions that Inform

Designing Communications that Highlight News Filter Bubbles
and Provide Strategies for Combatting their Negative Effects.

Min Kim (author)

Peter Scupelli (primary advisor)

Stacie Rohrbach (secondary advisor)

Abstract

As machine learning grows more sophisticated everyday, the societal impacts of algorithm models automating and dictating much of our online behaviors becomes undeniable. This has led to ethical problems in technology-based media (filter bubbles on facebook, search discriminations on google search, datafication and privacy in general, etc), and despite our growing dependency upon machine learning systems, we have no clear ethical guidelines to provide either the computer scientists when building the algorithms, or designers when implementing such systems. This thesis explores the friction points where design might intervene in ways that effectively address the challenges at hand, as well as better ways of designing users' relationships with their filter bubbles, through literature reviews, exploratory research, paper prototypes, and a survey. The goals are twofold: 1. To diversify news consumption practices, and 2. To encourage people to become more aware of their own behaviors on social media. After conducting literature reviews, I identified the following friction points: lack of transparency over what data is being pulled; and a general lack of user agency and control over the kinds of data they would like pulled from their engagement and the sorts of content they might like to be shown. As a basis for inquiry, this study questions if there's a way to leverage design to help people become more aware of their contributions to their own filter bubbles, instead of pushing people to engage with others who think differently. Findings from the survey suggest that people find value in the resulting prototype, but that the stakeholders would need to expand beyond social media users, in order to consider a financial incentive for the business.

Project Goals & Scope

In this chapter, I propose the project scope and address some key literatures that inspired the intent of this thesis.

Introduction

“Things mediate our perception of reality.”

- Edmund Husserl

Machine Learning is prevalent in all facets of our day-to-day lives, in one form or another. It's in Google Maps when it finds us the fastest way to get from Downtown LA to Santa Monica Beach; it's in the books that Amazon recommends to us (“The Life-Changing Magic of Tidying Up,” anyone?), the cat videos that YouTube suggests, and in the newsfeed that Facebook curates for us. It has become an integral part of our lives, and we depend on its promise of efficiency and customizability. It is served up in easy-to-consume bite sizes, and is always attractive. But how contextual is it? What if I want to read about how some people collect taxidermy with passion, instead of reading about the art of tidying up and living minimally? What if I want to stop seeing videos about cats with afros, and start seeing golden retriever videos for a change? What if I want to see my high school friends' feeds once in awhile, instead of those from friends who I see and talk to everyday? And what are the repercussions of being exposed to selective content that's been carefully curated for me?

By making assumptions about our habits and practices to make certain recommendations, algorithm models are often too quick to judge us and can therefore be disrespectful. This process is known as selective exposure, or confirmation bias. By funneling us down the paths of selective exposure, Filter Bubbles are created, as coined by Eli Pariser.

“Your filter bubble is your own personal, unique universe of information that you live in online. And what’s in your filter bubble depends on who you are, and it depends on what you do. But the thing is that you don’t decide what gets in. And more importantly, you don’t actually see what gets edited out.”¹

Algorithm models lead people to believe that certain products are the best they’ll find related to a particular task, or that a liberal satirist’s viewpoint—in the likes of Jon Stewart or Stephen Colbert— is the strongest and the most logical out there, disregarding the fact that there are a lot more liberal satirists than conservative ones,² and that the liberal audience is therefore less likely to hear from the Other Side (or any others).

The Association for Computing Machinery (ACM) presents us with a “Code of Ethics and Professional Conduct.”³ In an approach resembling the HIPAA Act for medical practitioners, ACM advocates for acting in the public interest, as opposed to the business’s: “to be accountable and responsible for their [business’s] work, to moderate private interests with public good, to ensure safety and privacy, to avoid deception, and to consider the disadvantaged,” and general moral imperatives of ACM include “avoiding harm to others”, “be fair and take action not to discriminate,” and “respect the privacy of others.”

1. Eli Pariser, “Beware Online ‘Filter Bubbles’” (TED Talk: 2011), https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles

2. Oliver Morrison, “Waiting for the Conservative Jon Stewart” (The Atlantic: 2015), <https://www.theatlantic.com/entertainment/archive/2015/02/why-theres-no-conservative-jon-stewart/385480/>

3. ACM Council, “ACM Code of Ethics and Professional Conduct” (Association for Computing Machinery: 2018), <https://ethics.acm.org/code-of-ethics>

In a similar vein, in his book "The Master Algorithm," Pedro Domingos states the importance of "fostering awareness [of what data's being gathered and the potential costs & benefits] and letting everyone make their individual choices about what to share, what not, and how and where."⁴

This idea of giving back control to the users brought to mind the 'Levels of Automation' chart developed by Parasuraman et al. They define automation as "the use of electronic or mechanical devices to replace human labor,"⁵ and in it, they identify ten varying levels of computer automation versus human-made decisions.

LEVELS OF AUTOMATION OF DECISION AND ACTION SELECTION

- HIGH**
10. The computer decides everything, acts autonomously, ignoring the human.
 9. informs the human only if it, the computer, decides to
 8. informs the human only if asked, or
 7. executes automatically, then necessarily informs the human, and
 6. allows the human a restricted time to veto before automatic execution, or
 5. executes that suggestion if the human approves, or
 4. suggests one alternative
 3. narrows the selection down to a few, or
 2. The computer offers a complete set of decision/action alternatives, or
- LOW**
1. The computer offers no assistance: human must take all decisions and actions.

Figure 1: Sheridan's Levels of Automation Chart.⁵

4. Pedro Domingos, *The Master Algorithm* (London: Penguin Books, 2015), 263-290.

5. Thomas B. Sheridan and Raja Parasuraman, "Human-Automation Interaction," *Reviews of Human Factors and Ergonomics* 1, no. 1 (2005)

The table (figure 1) starts off with level 1, computers offering no automation assistance, leaving all decisions and actions to the humans, gradually giving more authority to the computer at level 5 to execute one suggestion at human command, all the way to level 10, where the computer “decides everything, acts autonomously, ignoring the human.” In the case of a typical recommendation engine where content is curated at the backend without any knowledge of the user, the level of automation borders on a level 10. “Ignoring the human,” in that sense, becomes ignoring the context. Algorithms decide what to show and what not to show based on our past behaviors, but they don’t always show us what we ought to see; only the stuff that they deem will coax a click out of us.

How might we leverage design to re-humanize decision making processes, to help audiences make well-informed choices when interacting with automation systems and algorithm models?

That algorithm models make decisions discriminatory in nature and narrow our worldview is an unexpected byproduct of our wish to make modern digital lives more efficient, and it comes with unexpected politics attached. Algorithm models are an obvious example of a social determination system, which claims that “what matters is not technology itself, but the social or economic system in which it is embedded.”⁶ Winner argues also for the dangers of social determination in that it suggests that technical things do not matter at all. He then defines technological determinism: “unmediated by any other influences, [technology] molds society to fit its patterns.” In the case of algorithm models, it’s both socially deterministic—in that it amplifies what already exists in society—as well as a technologically deterministic system. Nonetheless, by forcing the models’ decisions upon people and affecting their lives without making the methods for such decisions explicit, it is also “molding the society to fit its patterns. In the following chapter, I will discuss the motivations behind tackling this very political, digital dilemma.

6. Langdon Winner, “Do Artifacts Have Politics?,” *Daedalus*, Vol. 109, No. 1, *Modern Technology: Problem or Opportunity?*, Winter 1980, The MIT Press.

Literature Review & Backcasting

In this chapter, I define the concept of selective exposure and its effect on our digital consumption practices and our society at large, and further explain the rationale behind this project. Then I discuss the key findings from literature review, as well as my interest in lightweight speculative fictioning and backcasting methods, which inform the direction of the thesis.

The effects of selective exposure

“Selective Exposure is a concept where rather than face the unpleasant feeling of having our beliefs tested, we simply steer clear of information that contradicts what we think we know. In other words, to avoid the cognitive dissonance that comes out of receiving news that challenges our beliefs, we cunningly select the messages we consume. This may be conscious or unconscious.”⁷

In his book *True Enough*, Farhad Manjoo speaks about the strength of the message being a key factor in how we decide to selectively expose ourselves to information: If we come across something Manjoo terms a “Strong Dissonant Information,” or “one that doesn’t support our views,” we avoid it. If we come across a “Weak Dissonant Message,” or “one that doesn’t support our views but is weak,” we may listen to it just because we want to refute it. If we come across a “Strong Consonant Message,” or “one that supports our views,” we consume it. If we come across a “Weak Consonant Message,” or messages that support our views but that we consider to be factually unsteady, we tend to ignore it.⁸

Practices of selective exposure, or confirmation bias, is prevalent everywhere and is often difficult to avoid. But social media sites, their curation of contents, is effectively reinforcing this phenomena, trapping us in the bubble that we inhabit much too comfortably.

To paint a picture of a possible consequence of this phenomena, I’d like you to think back to November 8th, 2016. Americans were waking up to the fact that the U.S. is much more divided in ideologies than some

7. “Selective Exposure,” *True Enough* by Farhad Manjoo, <http://trueenough.weebly.com/selective-exposure.html>

8. Farhad Manjoo, *True Enough: Learning to live in a post-fact society* (Hoboken, NJ: Wiley, 2008), 43.

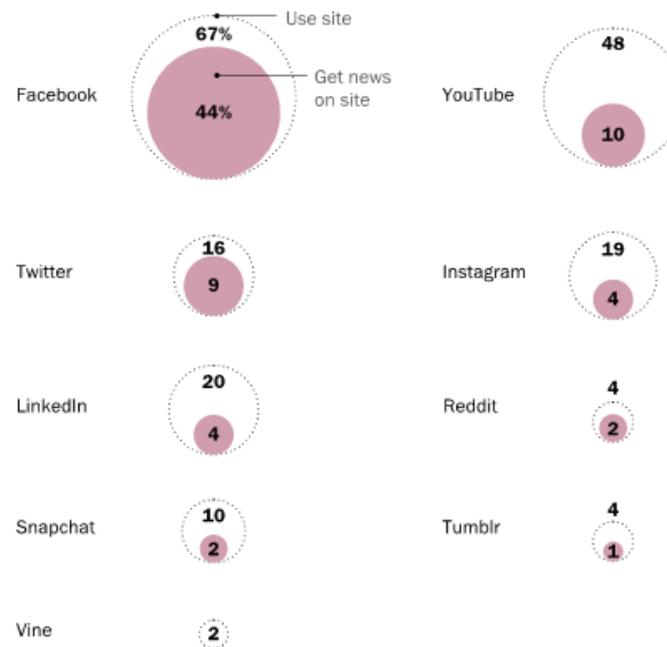
privileged enough to be called “liberal elites” realized, and that the democratic progress we’d been seeing in recent years was about to be stalled. Had we realized that the think pieces that we’d been reading in liberal publications like The Atlantic or CNN were mere reinforcements of the opinions of the minority and not the representation of the American voice that some of us thought that they were, would we have taken different measures? I am by no means trying to attribute the turnout of this recent election to filter bubbles; but would recognizing our tendencies to selectively expose ourselves have helped us bust out of the filter bubbles and brought us closer to understanding others’ perspectives, instead of personifying all of the ideas distant from our views into a laughingstock called Trump? Would one side have reached out to, read about, and spoken to those representing different perspectives, and promoted a solution that we could all peacefully agree on, instead of pointing fingers to laugh, vilify, and

to ridicule? According to this study by Pew Research Center, more than half of U.S. adults get news from social media platforms, Facebook “leading the pack.” (figure 2) ⁹

This is precisely the problem we are facing in the age where social media plays such a big part of how we digest the world’s information. Reading the news on Reddit, Twitter, or Facebook wouldn’t be such a big issue if we didn’t already tend to make new friends with those who already share our ideologies. Facebook is a culmination of humans wanting to feel as if they all share the

Social media news use: Facebook leads the pack

% of U.S. adults who ...



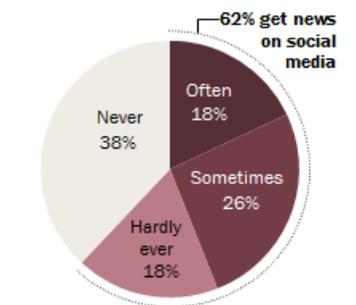
Source: Survey conducted Jan. 12-Feb. 8, 2016. "News Use Across Social Media Platforms 2016"

PEW RESEARCH CENTER

Figure 2: Over half of American adults get their news from social media, Facebook “leading the pack.”

About 6-in-10 Americans get news from social media

% of U.S. adults who get news on a social networking site ...



Source: Survey conducted Jan. 12-Feb. 8, 2016.

"News Use Across Social Media Platforms 2016"

PEW RESEARCH CENTER

9. Jeffrey Gottfried and Elisa Shearer, “News Use Across Social Media Platforms 2016,” Pew Research Center’s Journalism Project, May 26, 2016, <http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/>.

same political, religious, social, or just philosophical ideals. It offers the ease of sharing thoughts and learning others' by grouping one like-minded individual or group to another, but the price of its practices is that we're often not exposed to ideas outside of those bubbles. It is a paradigm shift from how we used to consume news, and it "raises questions about the ability of computers to curate news, a role traditionally played by [human] editors." ¹⁰

Selective exposure is especially problematic because Facebook curates the world's news for us, but always seems to avoid taking on such an important journalistic role. In an interview with the New York Times, an engineer at Facebook stated: "We try to explicitly view ourselves as not editors... We don't want to have editorial judgment over the content that's in your feed. You've made your friends, you've connected to the pages that you want to connect to and you're the best decider for the things that you care about." ¹¹

In their paper "Algorithmic Transparency in the News Media," Nicholas Diakopoulos and Michael Koliska talk about a certain journalistic responsibility that algorithmic systems have in the production of news: "Since algorithmic systems rely on quantification of the world in order to operate, human processes are needed to set the ground rules, definitions, and boundaries of that quantification in order to enable algorithmic operation at scale... [we] need to consider the human influences on that data." ¹²

10. Ravi Somaiya, "How Facebook Is Changing the Way Its Users Consume Journalism," The New York Times, October 26, 2014, <https://www.nytimes.com/2014/10/27/business/media/how-facebook-is-changing-the-way-its-users-consume-journalism.html>.

11. Ibid.

12. Nicholas Diakopoulos and Michael Koliska, "Algorithmic Transparency in the News Media," Digital Journalism (2016): 10

Literature review findings

Currently, Facebook's newsfeed algorithm has two main design goals:

1. To show everyone the right content at the right time so they don't miss the stories that are important to them.
2. To have newsfeed display posts more prominently that will generate more interaction or engagement.¹³

Both of these goals have led Facebook to successfully engage many like-minded people with each other, but encouraging an accurate and unbiased news digest isn't a priority. I propose that any social media site add another goal: to diversify news content to help make users become wise news consumers, and to empower them to be more knowledgeable about the world (not to mention, it would also add to the user delight at finding ideas that are novel to them). Because no matter what Facebook (or any other social media platforms) describe as their roles and goals in contacting people and delivering information, a significant amount of the world's population reads the news through these machine learning sites. Thus, they have a journalistic responsibility to their readers.

Diakopoulos and Koliska define transparency as "the ways in which people, both inside and external to journalism, are given a chance to **monitor, check, criticize** and even **intervene** in the journalistic process." It allows users access to more information, "which can influence power relationships between governments and citizens, business and customers, news outlets and audiences."¹⁴

13. "News Feed FYI: A Window Into News Feed," Facebook for Business, <https://www.facebook.com/business/news/News-Feed-FYI-A-Window-Into-News-Feed>.

14. Diakopoulos and Koliska, 3.

How might design help in better mediating the interaction between the machine learning algorithms and the decision making step, and provide users with transparency into the workings of the black box system that is social media, and therefore empower them with the means to **“monitor, check, criticize, and intervene”**?

If the main problem with the current way of digesting news has to do with lack of transparency—both in how social media platforms observe a user’s data and curate the content, as well as in how cognizant the users are of their own behaviors—how might we use design methods to test and shift such practices?

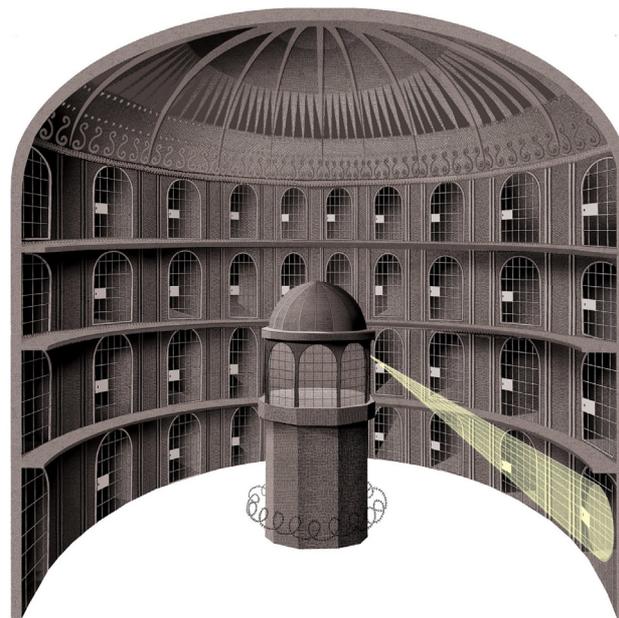


Figure 3. “The Panopticon,”
sketched by British philosopher
Jeremy Bentham.

In his thesis, “The design of our own lives Technical mediation and subjectivation after Foucault,” Steven Dorrestjin talks about a structure called The Panopticon, through the eyes of French social theorist Michel Foucault. The Panopticon is a fictional circular architecture developed by Bentham with a watchtower sitting at its center.¹⁵ The idea is that by constructing a watchtower in the middle, its inhabitants are always under the impression that they are being surveilled, and that this will lead to a more disciplined society overall.

15. “The Panopticon.” UCL Bentham Project, <https://www.ucl.ac.uk/Bentham-Project/who/panopticon>.

In the digital world, social media users recognize that data is being collected, but they do not pay attention to it because they can't see who is surveying them, what type of data is being mined, or how it's being used to curate predetermined experiences. Following Bentham's idea about self-surveillance leading to a more disciplined society, I am hypothesizing that we could start tackling the problem of passive news consumption practices by letting users observe their own behavior on social media, which would help them become more mindful of their news consumption practices. In sum, the main findings from my literature review are that:

1. There is a Need for Transparency

In his study, "Algorithmic Transparency in the News Media," Diakopoulos and Koliska claim that, because social media is increasingly becoming a place for news consumption (as proven by the Pew Research Center study), there's a "journalistic responsibility" to make the algorithm more transparent and to encourage trust. ¹⁶

2. Fairness should be reached through Awareness

In his book *Master Algorithm*, Pedro Domingos describes the value of "fostering awareness [of what data's being gathered and the potential costs & benefits] and letting everyone make their individual choices about what to share, what not, how, and where [to share]." ¹⁷

3. Tapping into the power of Self-Monitoring for a More Balanced Society

Bentham's Panopticon is structured so that it enables everyone to surveil themselves, making for a more disciplined and balanced society. Could the idea of self-surveillance help people become more conscious of their rather passive consumption habits online? (Dorrestijn, 1977) ¹⁸

16. Diakopoulos and Koliska.

17. Domingos, 268.

18. Steven Dorrestijn, "Technical Mediation and Subjectivation: Tracing and Extending Foucault's Philosophy of Technology," *Philosophy & Technology* 25, no. 2 (2011).

Lightweight speculative fictioning & backcasting

During my initial research phase, I had the opportunity to visit a friend's exhibition that explored the concept of speculative fiction in the areas of coloniality, environmental sustainability, and crises of culture and race climate change in both the Global South and North, titled "Climactic: Post-Normal Design." The exhibition's curators Ahmed Ansari and Deepa Butoliya, among others, presented "alternative models for design that broaden human capacity to understand and intervene in accelerated social and environmental crises," to think about the "ongoing moment of contingency and precarity that characterize contemporary life and challenge the design disciplines."¹⁹

Even though most pieces were mainly around the topic of fictional decolonial societies and environmental crises, the exhibition was provocative, and the conversations that followed on the consequences of current developments around the world involving design, even more so. I started thinking about applying speculative thinking in the context of social media, and I approached the exhibition's provocateur, Cameron Tonkinwise, to get his thoughts. As a result, some extremely thought-provoking ideas were thrown around. Sprung from an article on The Guardian "What would happen if we asked people to switch feeds w others for 5 wks prior to election?,"²⁰ Cameron suggested that I speculate on four to five wildly unconventional scenarios that focused on future social media usage, and to ask: how will people reshape behaviors around social media, and what would be the consequences?

19. "Miller Gallery at Carnegie Mellon Universtiy." Back to Yes. <http://millergallery.cfa.cmu.edu/exhibitions/climacticpostnormal/>.

20. Olivia Solon, Sam Levin, and Julia Carrie Wong, "Bursting the Facebook bubble: we asked voters on the left and right to swap feeds," The Guardian, November 16, 2016, <https://www.theguardian.com/us-news/2016/nov/16/facebook-bias-bubble-us-election-conservative-liberal-news-feed>.

Provoked by this idea, I dug into the concept of Futuring and Backcasting in the design discipline. I looked primarily to the course syllabus that Peter had put together for his course "Introduction to Ddesign the Future,"²¹ where students are asked to "explore the forces that drive change in the future (i.e., social, economic, political, environmental, technological), and learn to align innovation strategically with the trajectories of those forces on long-time horizons." In it, Peter describes instances where people take Futures thinking and combine it with Design Thinking. It can culminate into a powerful toolset, where the creation of multiple different worlds and the speculation of outcomes can be useful to gauge the ideal state.

If visualizing an ideal state can play a powerful role in leading to a desired result, then I proposed that establishing an ideal state could be used as a metric to measure the success of my designs. Further, I wanted the designers of social media platforms to be equipped with the means to visualize such a state, as I hypothesized that it would help them make informed design decisions. Then, we can backcast on what the current steps should be, for both of the stakeholders—the social media platforms, and their users.

My goal for the outcome is twofold: To create a toolset by which a designer might ideate on the ethical implications of a design decision within the realm of machine-learning, algorithm-based sites; and to develop a prototype based on previous learnings to field test the effectiveness of such a toolset.

21. "Introduction to ddesign the future | Introduction to Ddesign the Future on WordPress.com," Introduction to Ddesign the Future, <https://ddesignthefuture.wordpress.com/>.

Exploratory Research & Theories of Change

In this chapter, I cover the exploratory research conducted to

1. Confirm the filter bubble phenomena, and
2. Dismantle the different elements in a Facebook post, then discuss the key workshop findings.

Together with the findings from literature reviews, these inform my three hypotheses (theories of change).

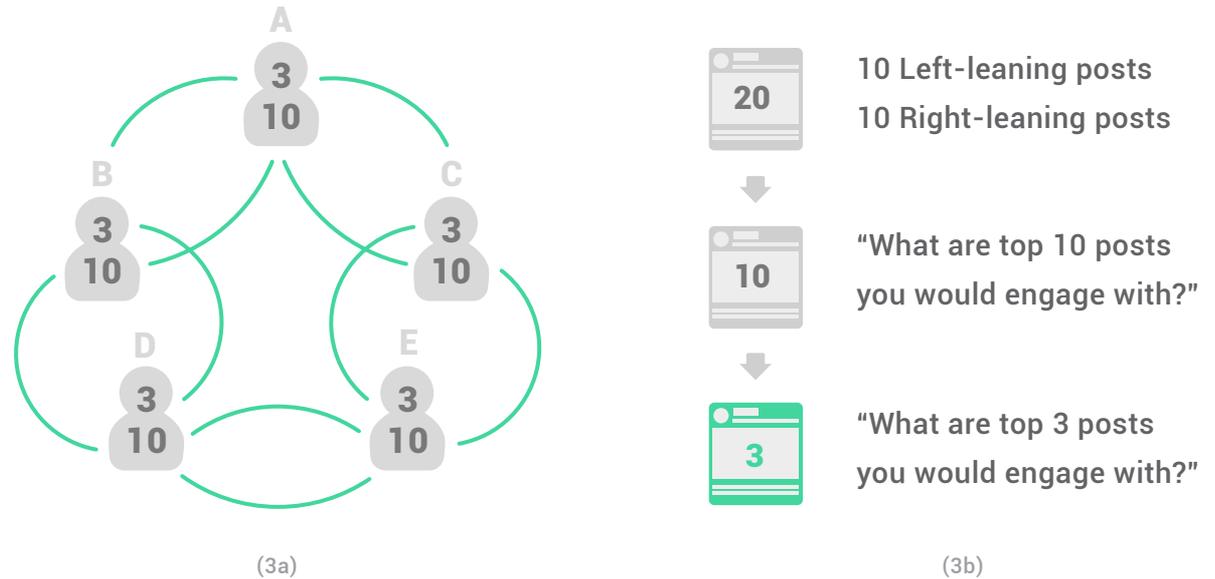
Exploratory research

During the exploratory research stage, I aimed to prove that people actually do perpetuate their own filter bubbles. After verifying my initial hypothesis, I built rapid prototypes to use to test my concepts.

With the help of PhD candidate, Ahmed Ansari (who is working on a similar agenda for his thesis, with different end goal and approach), I adopted a simplified research plan. The steps are as follows.

Step 1: Confirm the Filter Bubble Phenomena

Figure 3: Exploratory Research Methodology diagram. Step 1 is confirming the filter bubble effect, where each participant's top choices for posts they'd engage with would overlap with others of similar political leanings



Pictured above (figure 3a) is the rough plan to prove the Filter Bubble, and how people perpetuate it by clicking only on things that are familiar (A, B, C, D, E stand for the 5 participants).

1. Collect 20 newsfeed posts (mainly news articles), 10 left-leaning and 10 right-leaning
2. Present the 20 posts to 5 Left-leaning participants, and 5 Right-leaning participants
3. Ask them to choose 10 posts that grab their attention
From those 10, choose 3–5 posts that they would actually read (figure 3b).

Step 2: Dissect the Anatomy of a Post

As a means of discovering what specific elements affect the reader and their wish to engage, I took the same posts and separated them into four parts: 1. the name of the source (i.e. publication or a friend) and their commentary; 2. the photo of the article; 3. the headline and a short quote from the article; and 4. the bottom-most social engagement bar (likes, comments, number of shares). On a simple scale system, I labeled one end “Would engage” and the other “Wouldn’t engage.” I then asked the participants:

1. If you were to lay each of the elements out on this scale system, how would you rank them? Which element has the most impact on your decision to engage with the post?
2. How are you deciding what to engage with? What factors are being considered in your decision? (for a more qualitative data)

Findings

- Vague or extreme language used in article headline and introduction discouraged engagement.
- Majority of the participants ranked “friends sharing the post” as having the highest level of impact on their engagement.
- Study suggests that 67% of participants would actively contribute to their own filter bubbles.

Figure 4: Participant sorting different elements in Facebook news article posts: the photo, name of publication, article headline, article description, “shares” by friends, “likes” by friends, and “likes” by general public.

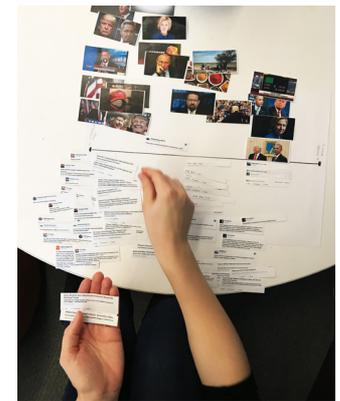
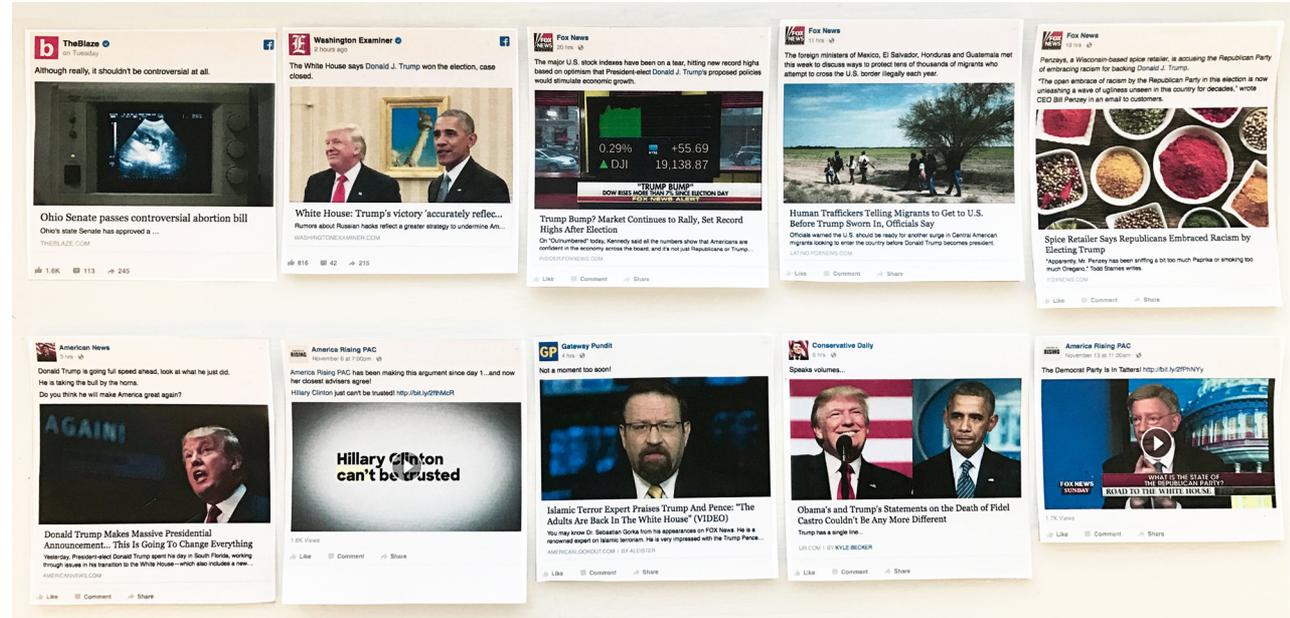


Figure 5 The 20 different Facebook article posts—one set written by left-leaning publications (5a), the other by right-leaning publications (5b)—used in exploratory research used for the initial research exercise.



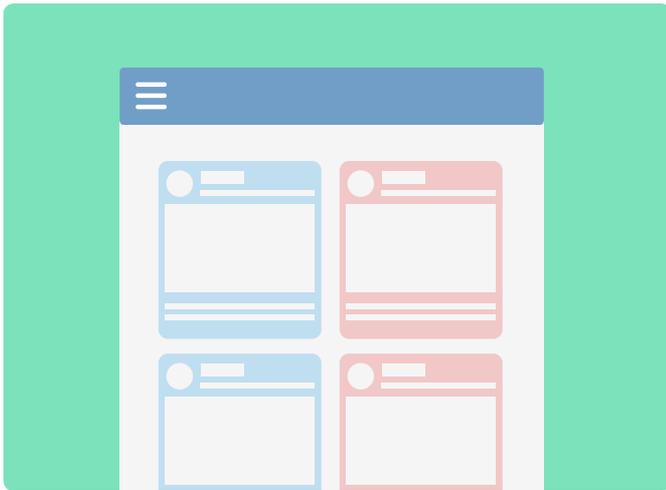
(5a)



(5b)

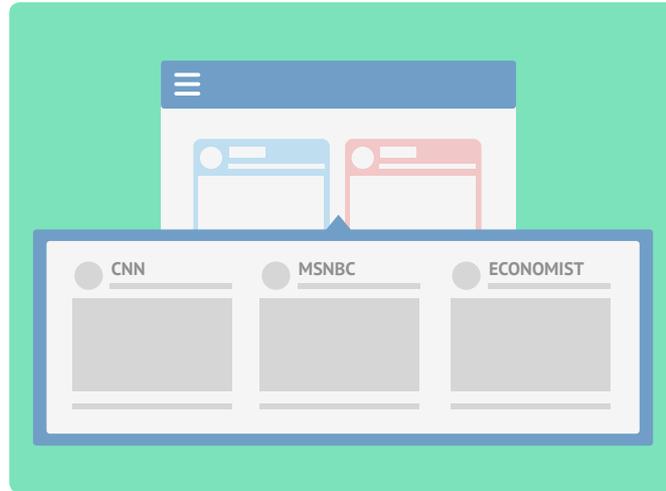
Theories of Change

The findings from the exploratory research combined with the three main literature reviews informed the three theories of change that emerged from several design interventions that I developed, to effectively further research *with and by design*.



Achieving Content Transparency for Better Post Exploration

During the exploratory research, one participant said that she “can’t really tell what the content of the post will be about [from the article headline or the photo], so I’ll pass.” Another said, “The headline indicates that it’s going to be too extreme and alt-right,” but yet another mentioned that “if an obscure publication has an interesting title, I’ll give it a skim.” To reduce the friction caused by the ambiguity in the posts, I am asking the question: How might the UI elements better aid in interpreting the obscurity of the content and facilitate a more neutral way of navigating through the news stories?



Mindful Engagement for Better Awareness

Can we design a way to help users become more aware of the effects of passive engagement in a machine learning system? How could we enable everyone to question their versions of reality?



Self-Monitoring for a More Balanced Media Exploration

Informed by the literature review on Bentham's Panopticon, the last theory of change is exploring the idea that instead of a top-down distribution of power, what would happen if we adopt a lateral one, where everyone becomes more conscious of their own actions? So far, the study suggests that a majority of participants actively contribute to their own filter bubbles. If we're reminded of this fact, would it motivate us to be more equally informed of ideologies from both sides of the spectrum?

Paper Prototypes & Findings

In this chapter, I cover the secondary research I conducted with paper prototypes that were developed from the three hypotheses, and discuss the findings that inform the final design intervention prototype.

Paper prototypes

After the initial rounds of research, I mocked up several very simple mid-fidelity designs using Photoshop, printed a blank Facebook newsfeed, several different news posts from CNN, FOX news, The New York Times, The Atlantic, and The Economist, each treating different topics of current events, and the four different mockups that were developed out of my three hypotheses. Each of the mockups addressed different aspects of the design problems that arose in the interviews and workshops. I then presented them as paper prototypes, and asked my participants to explore a printed facebook feed, which helped me learn which features garnered positive responses.

Feature 1: FAQ for Algorithm Model Transparency

This paper prototype features a small “FAQ” button to enable users to gain transparency into why they’re seeing a certain post. Upon clicking on the “You’re seeing this post because…” tab, users see the last three actions they conducted on Facebook that led its algorithm to show a particular post.

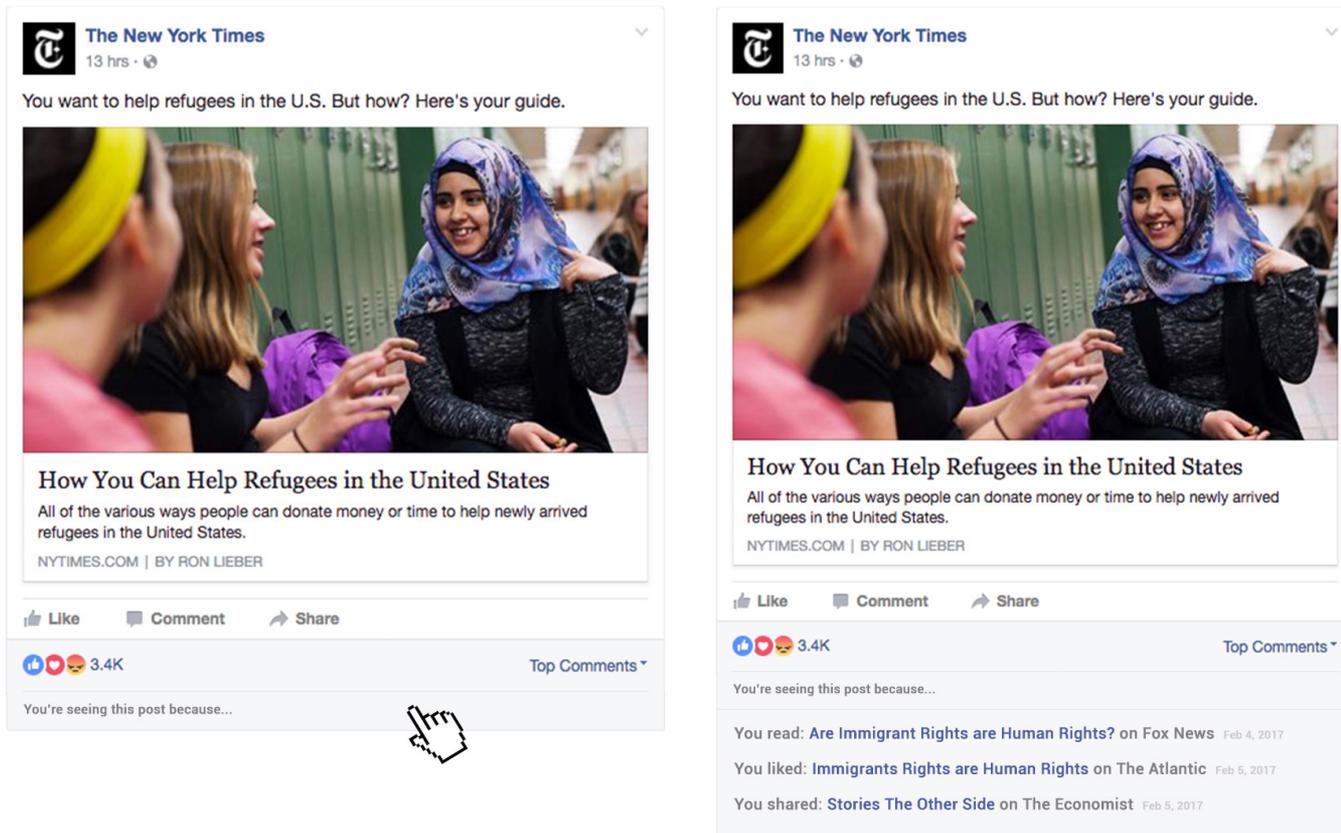
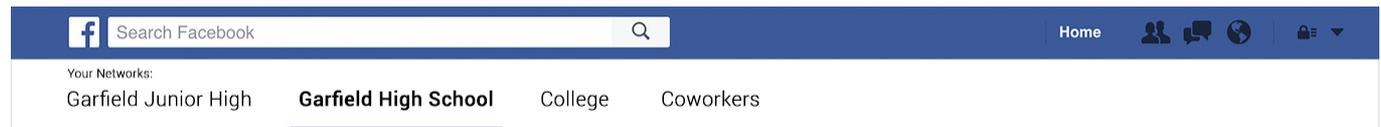


Figure 6: Feature 1's 'FAQ for Algorithm Model Transparency'.

Feature 2: Friend Groups' Feeds

Facebook keeps a track of various different kinds of metadata regarding its users in their "Ads preferences settings" page, usually quite hidden in the Privacy page. Here lies all the stuff that Facebook was able to deduce about you. It includes tags ranging from simple technical identifiers ("iPhone 7 user"), your late-night netflix inclinations ("Jiro Dreams of Sushi"), sociopolitical tags ("Very Liberal"), your ethnicity ("Asian-American" that is "Away from Home"), and much more that makes you go: "Okay, this is creepy. What else do they know about me?!" Using such metadata, this feature allows users to categorize their friend groups, be it by high school friends, college friends, or friends who are "politically conservative" as deemed by Facebook. It then lets you click through different types of feeds that your friend groups are seeing, on their newsfeeds, to give you a better idea of what kinds of content everyone else is seeing, and what types of "realities" each of us are living in.

Figure 7: feature 2's 'Friend Groups' Feeds'.



Feature 3: Hover Tool

With this feature, the user would hover or click a button with their mouse on a news article posted in their feed. After doing so, the user would be presented with similar news stories published by other perspectives. The hope is that the ease with which the users can apply this tool to their daily news feed consumption practices is respectful of their current method of usage, and yet the lightweight nature of the hover tool intervention will prove to be an effective way to provide a wider array of news to the users.



Figure 8: feature 3's 'Hover Tool'.

Feature 4: Meter Tool

Primarily based on Foucault's reading of Bentham's Panopticon, this meter tool is based on the third Theory of Change, where the inhabitants of Panopticon become more disciplined under the surveillance of the central watch tower. With each new piece of news content that the user engages with on Facebook, the meter keeps track of the consumption activities.

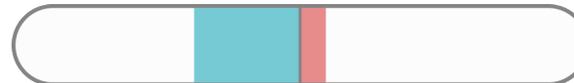


Figure 9: feature 3's 'Meter Tool'.

Participant #1 On feature #2: Friend Groups' Feeds

"I don't understand why I should care about what my middle school friends think or see? But I might be interested in seeing my college buddies' feeds, or my coworkers; it's more current and relevant [to me]."

Design opportunity: How might we ease the process of customizing friend groups?

Participant #2 Response on feature #2: Friend Groups' Feeds

"Facebook already categorizes your friends for you, but being able to click through certain friend groups' feeds and not their posts might be more interesting. But then again, I'm perfectly happy in my own bubble."

Design opportunity: How might we design the feed feature to be more engaging and prominent? Would it become more prevalent in people's FB feed practices?

Participant #2 Response on feature #3: Hover Tool

"If you make it past facebook, and target people that really care about news, it'd be useful."

Participant #3 Response on feature #1:FAQ for Algorithm Model Transparency

"Sometimes [on Facebook], I see the same things multiple times over. Sometimes I want to know why Facebook is so relentless about recommending some posts more than others."

The Design Intervention

In this chapter, I go over the chart that I've developed as a compilation of all the research findings. Then, I decide on the final prototype direction from paper prototype workshops, and add a layer of legitimacy to the meter tool, backed up by research. Finally, I cover the survey format, the results, and the limitations.

The Chart to provide an ethical design foundation

Based on literature reviews, interviews, and couple of rounds of exploratory research, I discovered numerous useful findings that could inform design concepts for the next big social media platform. I came to realize that no single platform can address transparency and awareness issues in one perfect solution. Thus, I started developing a set of recommendations for future designers who would design for Social Media platforms (figure 10). This is also an attempt to apply some of the backcasting methods I've repurposed from speculative fictioning and Futuring, where a designer equipped with this chart would work towards meeting the goal of the "Ideal State" in the last column. I will walk through the purpose and intent of each column in this chart.

Figure 10: First draft of the Chart

Categories	Exploratory Research Findings (Match up findings to the "Principles")	Guiding Principles (from lit reviews)	Generative Prompts	Speculative Consideration (of potential impacts)	Ideal State
Transparency		Transparency of whether or not a certain interaction is being recorded by the system / feeding into their algorithm model of you	How about an on/off switch? or a FAQ under each post?	Would it help users combat filter bubbles?	Users know why they're seeing what, and navigate consciously.
		The posts should be less obscure about the nature of their content	Can the post's visual hierarchy more accurately reflect the significance of the different elements in a post? (i.e. photo is not always representative of content, yet it takes up most of height)	Will there be a need to minimize visual discord to alleviate confusion? OR will there be a learning curve?	Design of post better indicates the nature and voice of the content
Awareness		It should be clear what inferences the systems are making about the user, how & when (i.e. based on the user's behavior)	"Algorithms at Work" badge to indicate that the system is actively churning to learn more about user (Diskopoulos, Koliska, 9)	How much self-regulation/ monitoring would a user be interested in engaging? How can design be leveraged to incentivize this behavior?	It is clear to distinguish what behaviors are contributing to a resulting feed
		People often grow out of an opinion or practice, and the ML systems don't catch on fast enough	Can there be different models of one user, that the user can toggle between, depending on needs?	Would any one 'model' be rich enough to represent the user well? Would they be too fragmented to be conveniently used universally?	Users have more control over what viewpoints they might be shown
		It is a journalistic responsibility for a news aggregator to provide various different content (Users often don't know how different their feeds or the news contents are from others)	How about a light-weight feature or tool that easily presents various different views?	Would it make users more active participants when reading a variety of news, as opposed to a passive one?	Users acknowledge that others might read different kinds of news (Helps users diversify their news digest practice), and are motivated to read up on how others think.
		*** bold: addressed through research & prototype & survey			

I chose two categories from literature reviews that I deemed important with which to work: Transparency and Awareness. These two subjects came up repeatedly in my literature study, and they proved to be the most contentious subjects with the study participants.

The second column titled “Exploratory Research Findings” is for designers to fill out as a follow-up to their own initial exploratory findings. Ideally, the findings that go into this section would correlate with the content in the rest of the row and across all columns, to ultimately match the “Ideal State” column to the far right.

Guiding Principles (from lit reviews)	Generative Prompts	Speculative Consideration (of potential impacts)	Ideal State
Transparency of whether or not a certain interaction is being recorded by the system / feeding into their algorithm model of you	How about an on/off switch? or a FAQ under each post?	Would it help users combat filter bubbles?	Users know why they're seeing what, and navigate consciously.
The posts should be less obscure about the nature of their content	Can the post's visual hierarchy more accurately reflect the significance of the different elements in a post? (i.e. photo is not always representative of content, yet it takes up most of height)	Will there be a need to minimize visual discord to alleviate confusion? OR will there be a learning curve?	Design of post better indicates the nature and voice of the content
It should be clear what inferences the systems are making about the user, how & when (i.e. based on the user's behavior)	“Algorithms at Work” badge to indicate that the system is actively churning to learn more about user (Diskopoulos, Koliska, 9)	How much self-regulation/ monitoring would a user be interested in engaging? How can design be leveraged to incentivize this behavior?	It is clear to distinguish what behaviors are contributing to a resulting feed
People often grow out of an opinion or practice, and the ML systems don't catch on fast enough	Can there be different models of one user, that the user can toggle between, depending on needs?	Would any one 'model' be rich enough to represent the user well? Would they be too fragmented to be conveniently used universally?	Users have more control over what viewpoints they might be shown
It is a journalistic responsibility for a news aggregator to provide various different content (Users often don't know how different their feeds or the news contents are from others)	How about a light-weight feature or tool that easily presents various different views?	Would it make users more active participants when reading a variety of news, as opposed to a passive one?	Users acknowledge that others might read different kinds of news (Helps users diversify their news digest practice), and are motivated to read up on how others think.

Figure 11: The chart in detail.

The next column, "Guiding Principles," is a collection of strongly recommended design decisions, derived from the literature review findings (figure 11). The next column to the right, "Generative Prompts," provides a pedestal to help designers start generating ideas.

The second to last column, "Speculative Considerations," aims to enable designers to consider the hypothetical yet probable implications of a design decision.

Last but not least, the final column lists the ideal states of all the facets in the five different rows of design considerations, to serve as the metrics of success, which weigh the successes and failures of each new design intervention in a social media platform.

With this particular prototype, I chose the last row as my focus. The ideal state here is:

Users acknowledge that others might read different kinds of news, and are motivated to read up on how others think," to help users diversify their news digest practices.

Some feedback I received from Stacie on the first draft of the chart involved:

Provide more context to the purpose of the chart by answering questions like:

- Who is it for? What is the purpose of each item? What do you hope the designers would get out of it? What kinds of challenges would you like for them to tackle, and why?

Give rationale to each of the findings and action items

- "You should consider this factor when designing, because it will afford transparency for the context of individual posts."

Clearly distinguish Recommendations that set a guideline and strongly suggest a certain practice, **versus actionable questions** that prompt the designers to ideate.

- Recommendation: "This practice is strongly encouraged, if you want to encourage awareness of their own online behaviors."
Actionable Question: "How might you do this and that?"

Set up the whole chart as **steps that a designer can take to reach an Ideal State.**

Developing the Prototype

Participant responses from the paper prototypes provided interesting insights into what people wanted out of a social media intervention tool. They expressed a strong interest in a lightweight tool that offers them diversity in news content. Based on this finding, I further developed the hover tool and the meter tool.

The hover tool was received favorably, because it offered an easy, painless way to read various perspectives on a single contentious topic. Participant responses for the meter tool were also interesting in that they pointed to a tendency to maintain a comfortable level of ignorance. Participants mentioned that they were “comfortable living in my filter bubble,” or that “being reminded that I live in a bubble will make me feel uncomfortable.” Another participant posed a very useful question: “Who’s determining whether the article is biased one way or the other? I feel judged.”

This last finding informed the decision to add a layer of legitimacy. I took the Pew research titled “Media Outlets by the Ideological Composition of Their Audience”²² and gave each news source a “score,” ranging from -10 (minus ten) for the audience base being extremely liberal, to 0 for a neutral audience base, to +10 (plus ten) for an extremely conservative audience base. Based on of this system, the meter would display the appropriate scale indicating the level of the user’s conception practice.

22. Benjamin Wormald, “Ideological Placement of Each Source’s Audience,” Pew Research Center’s Journalism Project, October 20, 2014, http://www.journalism.org/2014/10/21/political-polarization-media-habits/pj_14-10-21_mediapolarization-08/.

Ideological Placement of Each Source's Audience

Average ideological placement on a 10-point scale of ideological consistency of those who got news from each source in the past week...

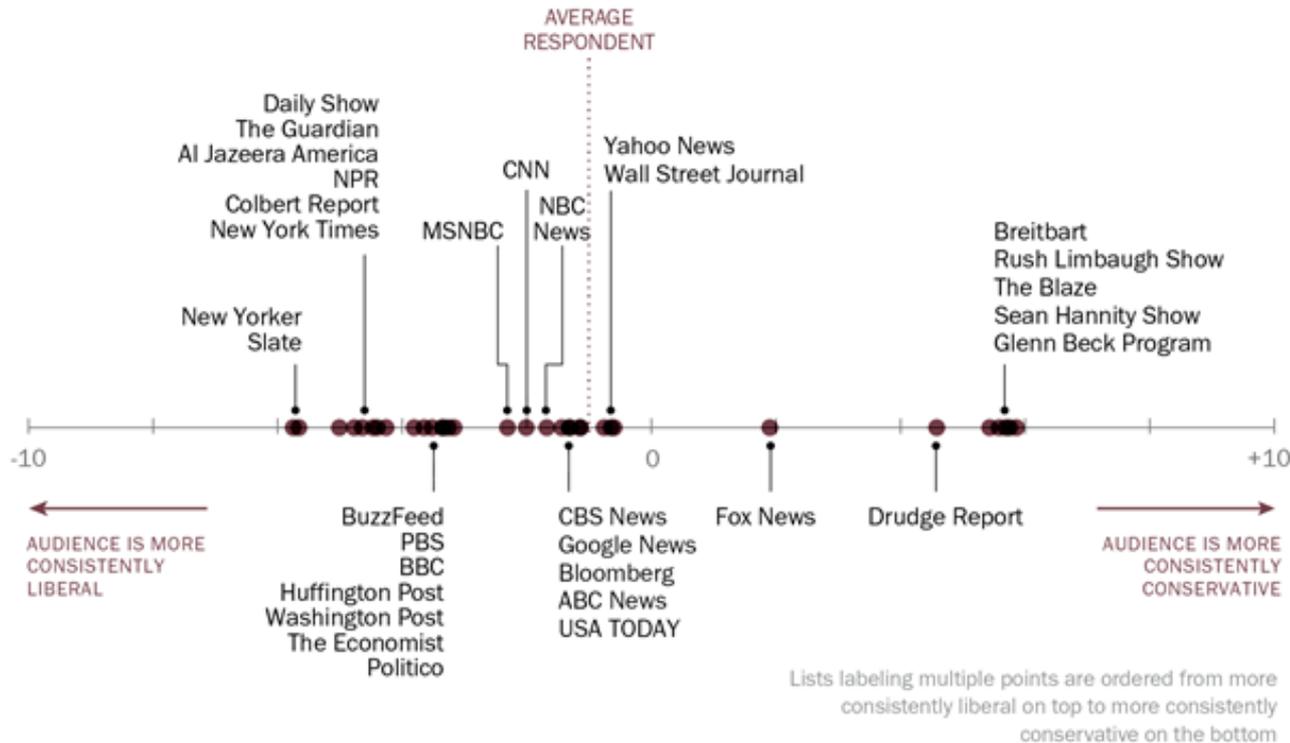


Figure 12: Pew Research Center's graphic scale for "Ideological Profile of Each Source's Audience [political leanings]"

American Trends Panel (wave 1). Survey conducted March 19-April 29, 2014. Q22. Based on all web respondents. Ideological consistency based on a scale of 10 political values questions (see About the Survey for more details.) ThinkProgress, DailyKos, Mother Jones, and The Ed Schultz Show are not included in this graphic because audience sample sizes are too small to analyze.

PEW RESEARCH CENTER

Survey Format, the results and limitations

With the developed prototypes, I developed and sent out a Qualtrics survey to gauge acceptance of the prototype hover tool, and to validate the helpfulness of the guidelines chart, to see whether or not the Ideal State on my chart had been achieved.

The survey was distributed via Facebook wall, closed Facebook groups, and personal email lists. Its format comprised of 1. Questions on demographics, 2. General Facebook usage practice, 3. Questions to help gauge responses to the Prototypes, and 4. Their Political leanings.

The survey remained open for approximately two weeks, and the results were divided but mostly positive. It's important to note that since a large majority of the participants were between 19 to 30 years of age and mostly liberal, the results probably don't accurately reflect the local population. On average, the four-page survey took participants approximately 10 - 15 minutes to complete, inclusive of the two-minute prototype concept video.

Q 22.

Imagine that you've read The Atlantic's refugee article shown above and that the hover tool provided you with opportunities to explore other perspectives.

Would you click on any of the other articles?

Out of 54 responses to the question above that directly targets the hover tool (figure 13), 50% of the participants said that they would click on an article with perspectives different than their own, and 30% said that they might. From these findings, I inferred that the hover tool would prove to be effective in helping users gain perspectives that differ from their own.

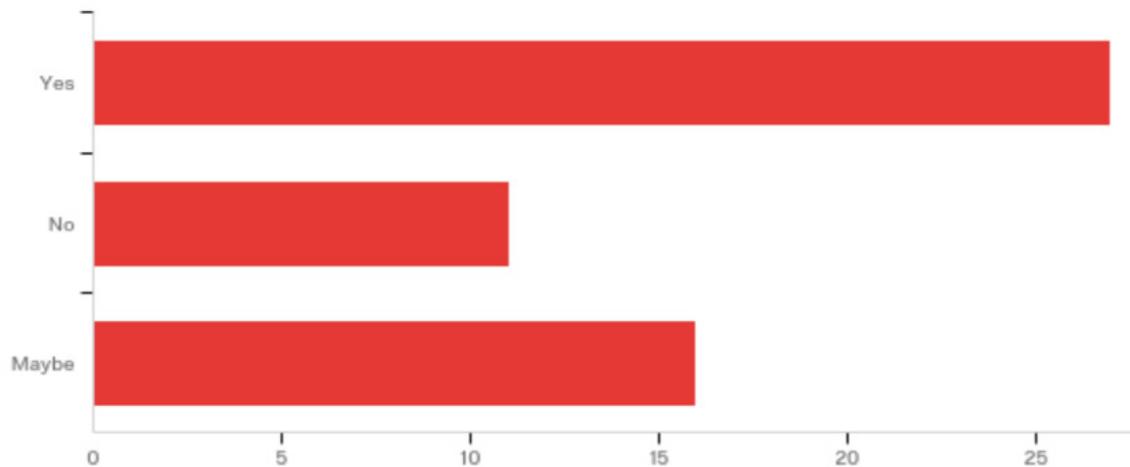


Figure 13: Survey results from Question #22.

Figure 14: Screenshot from survey question #29 regarding the meter tool

When reviewing the video, did you find the tracking meter (pictured below) helpful or annoying? Please explain.



In response to the question above (figure 14) that directly targets the usefulness of the meter tool, participants' responses were more disparate than the previous question. Some said, "Yes, I'd be more cautious [about my own bubble-reinforcing online habits]. But another participant said, "No [I wouldn't find the meter tool helpful], it's annoying," and yet another said, "No, it's just too much information."

Q 37.

**Please think of someone in your family or friends that might benefit from this tool.
What is their relation to you and how do you think this tool would benefit them?**

The next question (Q37) prompted the participant to think of a family member or a friend who they would want to recommend this to. Interestingly, this question was received much more positively than the previous one: *"My parents are so blinded by bias, they often believe in false information." It was ironic that only 50% wanted to use tool, 30% maybe, but almost all participants had someone to recommend the tool to. I thought that this discrepancy in the participants' responses might prove to be a fruitful design opportunity to pursue in the future.*

Future Implications

I started this project because the use of machine learning in systems—small and large—excited me. However, I was further motivated to investigate the topic when I learned that there was a lack of effective ethical toolkits that served as good reference and application guides for designers or researchers. The possibilities are limitless, but so are the consequences. I thought that by having a roadmap that compiles some of the main publications that focus on the usage and practices of machine learning, social media, and journalism, weighed against the receptivity of their users, we might be able to promote effective, active news consumption practices, and aid in curtailing some of the problematic consequences.

As mentioned in the previous chapter, the results of this survey aren't representative of the local population, therefore more accurate sampling of the political environment of social media needs to be executed and applied to the tool in order for it to be deemed accurate.

Nevertheless, here are some aspects of the tool that I believe hold potential in aiding news consumption and point to future research opportunities.

1. The hover tool, in particular, could be used to embed several functionalities, like filtering by different groups of people for their unique voices (e.g. "I want to hear from Women of Color").
2. As much as I'd like to enable users to transition from passive to active news consumers, the most ideal scenario would be for users to gain exposure to, and further seek out on their own, the actual facts and figures, untainted by the inherent bias brush of the media. A first step in this process would be to give them the means to explore different views, but at the mercy of journalists. The second step would expose them to important facts and figures, so that they could judge for themselves their stances without bias swinging them left and right.
3. The meter's simplicity in categorizing and displaying binary results (left to right, blue to red) might be a source for misinformation, and does not represent users who don't always identify with either, both, or no particular sides. Reminiscent of how the quantified the self is often void of human nuances, a future iteration of the binary meter may need to account for such subtleties, by either being more ambiguous in representing one's leanings, or more accurately representing the political landscape (figure 16) ²³ than is shown in its current form.

23. Dan M. Kahan, "Cultural Cognition as a Conception of the Cultural Theory of Risk," Handbook of Risk Theory, 2012.

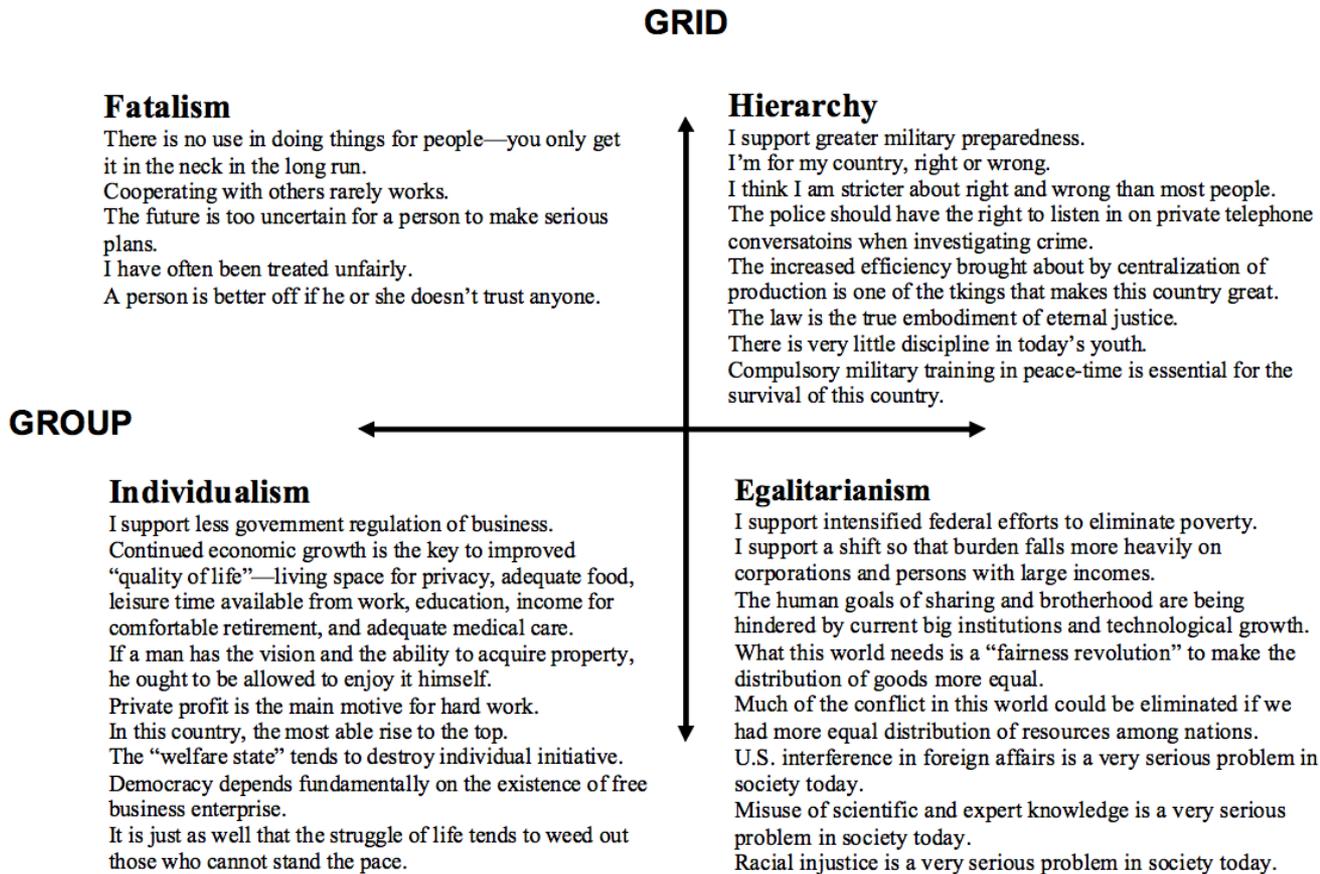


Figure 16: Shown here is Dan Kahan's Cultural scales matrix, representing different cultural, social, economical, and political ideologies, on "grid" and "group" axes. Each of the four identifier labels come with example scenarios to help better identify oneself.

4. Currently, this prototype is built as a third party plug-in, but the end goal is to integrate it into the main platform, which would mean making many assumptions regarding Facebook's motivation to integrate it. Some questions that need to be answered are:

- How much of Facebook's own algorithms should interact with the tool?
- What would be their financial motivation to be impartial by showing them neutral results, and to integrate the tool?
- How might we convince a publicly traded company to change their way of doing business to be more ethically conscious of machine learning and algorithmic practices?
- What would be the risks of changing their way of operating?
- Would ultra-conservative groups boycott Facebook if they showed different perspectives?

If we are to advocate for a more seamless collaboration between businesses, their users, and/or third-party vendors, we might need to tackle their business plans as one of the first steps to take.

5. This prototype, as part of an Human-Computer Interaction course capstone project, was developed into a functioning Chrome Extension (dubbed "ChromeView"), and is now available for download in the Chrome web store. Although it was developed too late into the year to allow time for a full research, it would be interesting to run fly-on-the-wall research sessions, or even extensive diary studies, where I would ask participants to write their thoughts after experiencing the prototype in their actual Facebook feeds for 20 minutes everyday. I imagine the results would be very different and much more nuanced, as they wouldn't need to predict their behaviors based on hypotheticals anymore.

References

Pariser, Eli. "Beware online 'filter bubbles'." | TED Talk | TED.com. https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles.

Morrison, Oliver. "Waiting for the Conservative Jon Stewart." The Atlantic. February 14, 2015. <https://www.theatlantic.com/entertainment/archive/2015/02/why-theres-no-conservative-jon-stewart/385480/>.

"ACM Code of Ethics and Professional Conduct." Association for Computing Machinery. <https://www.acm.org/about-acm/acm-code-of-ethics-and-professional-conduct>.

Domingos, Pedro. *The master algorithm: how the quest for the ultimate learning machine will remake our world*, 263-290. London: Penguin Books, 2015.

Sheridan, Thomas B., and Raja Parasuraman. "Human-Automation Interaction." *Reviews of Human Factors and Ergonomics* 1, no. 1 (2005): 89-129.

Winner, Langdon. "Do Artifacts Have Politics?" *Daedalus*, Vol. 109, No. 1, *Modern Technology: Problem or Opportunity?*, Winter 1980, 121-36. The MIT Press.

"Selective Exposure." True Enough: Learning to live in a post-fact society by Farhad Manjoo. <http://trueenough.weebly.com/selective-exposure.html>.

Gottfried, Jeffrey, and Elisa Shearer. "News Use Across Social Media Platforms 2016." Pew Research Center's Journalism Project. May 26, 2016. <http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/>.

Somaiya, Ravi. "How Facebook Is Changing the Way Its Users Consume Journalism." The New York Times. October 26, 2014. <https://www.nytimes.com/2014/10/27/business/media/how-facebook-is-changing-the-way-its-users-consume-journalism.html>.

Diakopoulos, Nicholas, and Michael Koliska. "Algorithmic Transparency in the News Media." *Digital Journalism*, 2016, 1-20. doi:10.1080/21670811.2016.1208053.

"News Feed FYI: A Window Into News Feed." Facebook for Business. <https://www.facebook.com/business/news/News-Feed-FYI-A-Window-Into-News-Feed>.

"The Panopticon." UCL Bentham Project, <https://www.ucl.ac.uk/Bentham-Project/who/panopticon>, accessed Nov. 17, 2015

Dorrestijn, Steven. "Technical Mediation and Subjectivation: Tracing and Extending Foucault's Philosophy of Technology." *Philosophy & Technology* 25, no. 2 (2011): 221-41.

"Miller Gallery at Carnegie Mellon Universtiy." Back to Yes. <http://millergallery.cfa.cmu.edu/exhibitions/climacticpostnormal/>.

Solon, Olivia, Sam Levin, and Julia Carrie Wong. "Bursting the Facebook bubble: we asked voters on the left and right to swap feeds." The Guardian. November 16, 2016. <https://www.theguardian.com/us-news/2016/nov/16/facebook-bias-bubble-us-election-conservative-liberal-news-feed>.

"Introduction to ddesign the future | Introduction to Ddesign the Future on WordPress.com." Introduction to Ddesign the Future. <https://ddesignthefuture.wordpress.com/>.

Wormald, Benjamin. "Ideological Placement of Each Source's Audience." Pew Research Center's Journalism Project. October 20, 2014. http://www.journalism.org/2014/10/21/political-polarization-media-habits/pj_14-10-21_mediapolarization-08/.

Kahan, Dan M. "Cultural Cognition as a Conception of the Cultural Theory of Risk." *Handbook of Risk Theory*, 2012, 725-59.