

New Tab Page Recommendations Strongly Concentrate Web Browsing to Familiar Sources

Homanga Bharadhwaj
Computer Science, IIT Kanpur
homangab@cse.iitk.ac.in

Nisheeth Srivastava
Computer Science, IIT Kanpur
nsrivast@cse.iitk.ac.in

ABSTRACT

To what extent do contemporary web technologies that seek to personalize users' browsing experience trap them in filter bubbles? Existing research has yielded mixed results on the possibility of such influence, but empirical research in this area has been entirely cross-sectional. In this paper, we report results from a longitudinal controlled web experiment conducted to determine whether passive recommendations embedded in common computer user-interfaces (UIs) reinforce users' habitual browsing behaviors, to the detriment of the diversity of the set of pages they tend to visit online. Inspired by classical demonstrations of a part-set cueing effect in memory, our experimental design manipulates the behaviour of the 'New Tab' page for consenting volunteers over a two month-long period in randomized time blocks of equal length. Analysis of their browsing behavior shows that users visit on average 15% fewer unique web pages while their browser's 'New Tab' page displays recommendations based on conventional frequency and recency-based algorithms, than if the display is left blank. This effect is seen systematically for all participants in our study. Further analysis of browsing behavior in this experiment clearly identifies the source of the difference between these modes of browsing: users consistently visit a greater diversity of web pages while typing in URLs in the URL/search bar when there are no recommendations on the 'New Tab' page. Finally, using a simulation study, modelling user behavior as a random walk on a graph, we extracted quantitative predictions about the extent to which discovery of new sources of information may be hindered by the concentration of browsing driven by such personalized 'New Tab' recommendations in classical browser UIs.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; *User studies; Web-based interaction*;

KEYWORDS

filter bubble, personalization, habitual web browsing, information diet, cognitive science

ACM Reference Format:

Homanga Bharadhwaj and Nisheeth Srivastava. 2019. New Tab Page Recommendations Strongly Concentrate Web Browsing to Familiar Sources. In *11th ACM Conference on Web Science (WebSci '19), June 30–July 3, 2019, Boston, MA, USA*. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3292522.3326011>

1 INTRODUCTION

The medium is the message [19]. McLuhan's gnomish aphorism has, in the half-century since its utterance, launched countless humanistic appraisals of the manner in which the form of modern communication technologies have shaped, and continue to shape the way people think. Given the extent to which web-based technologies now saturate our lived environment, substantiating these critiques quantitatively is very important.

For instance, it seems intuitively plausible, even likely, that the hyper-linked nature of web interfaces, in contrast with the linear nature of books, leads to greater distraction and shallower processing of information [5]. A survey of recent empirical research addressing this question, however, notes that even though the basic hypothesis is borne out - students reading material on screens do worse on tests than students reading material on paper, this performance differential appears to reduce quite rapidly with practice [20]. While it is still too early to conclude definitively either way about the validity of this particular critique, a general point remains: access to empirical data permits broad critiques of technology's influence on cognition to themselves be substantiated.

A similar story can be told about yet another intuitively appealing hypothesis - the filter bubble [24]. Here, the claim is that web-based recommendation services, by virtue of their algorithmic logic, recommend digital objects to people that they are already known to prefer, thus reducing the possibility for them to experience new content that they may or may not like. Again, the claim is easily believable, but has proved difficult to substantiate empirically. An analysis of the effect of accepting movie recommendations from the Movielens system showed that while the set of movies recommended became gradually less diverse with time, users who chose to follow the system's recommendations received more diverse recommendations than users who ignored them [22]. Recent research seeking to evaluate the existence of filter bubbles in Google News' personalization algorithm reported extremely minor differences in news articles presented to different user accounts explicitly set up to provoke differentiation [13]. Mixed evidence for filter bubbles was obtained from an analysis conducted by Microsoft on the news reading habits of 50000 Internet Explorer users, which found that while political articles accessed via social media or search engines are more politically extreme than those accessed directly from

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WebSci '19, June 30–July 3, 2019, Boston, MA, USA

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6202-3/19/06...\$15.00

<https://doi.org/10.1145/3292522.3326011>

websites, the use of these channels is also more associated with exposure to opposite political views than direct access [10].

There is an interesting dichotomy between theoretical and empirical evaluations of the filter bubble. Whereas theoretical critiques focus heavily on how recommendations could potentially narrow the diversity of users' demand patterns for digital media [24], empirical research focuses on identifying whether recommendation systems are narrowing the diversity of content providers' *supply* of digital objects [13, 22]. Work that explicitly tracks users' *consumption* patterns does appear to detect a small bubble effect [10], but the restriction of this study to studying only political diversity reduces the generalizability of its results.

In this paper, we report results from an experimental study of an extremely general version of the filter bubble hypothesis, evaluated strictly on consumption patterns rather than supply, and without restricting the nature of content consumed. Specifically, we investigate whether the presence of frequently visited tabs on 'new tab' pages in modern browsers leads to a concentration of browsing behavior into a more restricted set of websites than when such recommendation-based displays are absent. We conducted a controlled, longitudinal, within-subject experiment wherein we manipulated the behavior of consenting volunteers' new tab page displays without their knowledge and analyzed their browsing history at the end of the observation period to measure the degree to which our covert manipulations influenced users' web browsing behavior. We found strong statistical support for our hypothesis. Usage of new tab pages with personalized recommendations considerably reduced the number of unique websites visited by users, suggesting concentration of website browsing repertoires. Additional analyses revealed that this concentration was not uniform, and was seen most clearly precisely during browsing events most strongly associated with free memory recall - users typing URLs into address bars in the new tab page display, with recommendations visible underneath.

To assess the extent to which this measured concentration in web information sources could affect the diversity of users' experience on the web, we also ran a simulation study, using our data to set parameters for a random walk model of a web surfer seeking information from diverse sources. We found that, for empirically grounded values of non-habitual browsing, the size of suppression effects seen in our empirical sample would correspond to reductions in the frequency with which diverse information sources are accessed, of nearly 50% for the median user.

Thus, in this paper, we present empirical and simulation evidence to show that the presence of a set of web page icon recommendations on browser new tab pages reduces users' propensity to access infrequently visited web pages by nearly half. We also discuss the implications of this striking finding on the filter bubble literature, the principles of designing communication interfaces, and ethical considerations of web-mediated communication.

2 A CONTROLLED FILTER BUBBLE EXPERIMENT

We anticipated large individual differences in browsing behavior across subjects, suggesting that cohort-level differences may or

may not be representative of the median user's experience. Therefore, we focused on designing a within-subject experiment, with experimental manipulations that each participant in the experiment would experience in different time blocks.

2.1 Experiment design

We designed a Firefox web extension using JavaScript, HTML and CSS. It broadly consists of three parts, a webpage to replace the default new tab display, a browser action page to log the browsing data into a local file and a remote script located in the experimenter's server that is used to tweak the type of websites displayed in the new tab page, as shown in Figure 1. In this design, the experimenter does not have access to the participant's data during the course of the experiment. At the end of the experiment, the experimenter explicitly asks for the logs, and participants are advised to curate their history to remove personal and private content before handing them over, if they choose to do so.

We programmed four different behavior modes in the replacement new tab page generated when people install our extension, controlled by a parameter in the extension script that we could change remotely without the user's knowledge. Each participant's replacement new tab page shows the web page icons as per exactly one of the following four modes at a given time.

- (1) Display the **most** visited sites in browsing history
- (2) Display the **least** visited sites in browsing history (at least one visit, naturally)
- (3) **Default** behavior, following Mozilla's frequency algorithm - a combination of frequency and recency [11]
- (4) Display a **blank** page

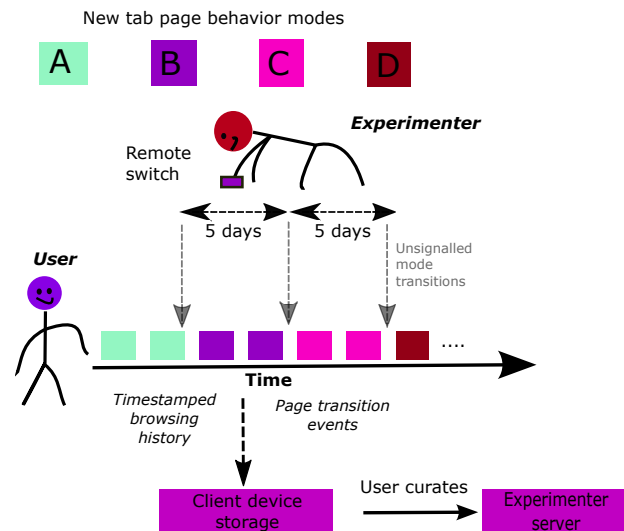


Figure 1: The experiment design

Our web-extension logs the webpages the user visits, the time of visits and the transition type from one webpage to the next - **typed** in the address or search bar, **linked** to the content of

a different site, or **clicked** through the browser UI. It also automatically logs all the websites currently displayed in the new tab page whenever it is opened. The replacement new tab page is designed to visually resemble the default Firefox new tab page so that after a certain period of time, the user is not very conscious of the extension running in background. Once the extension is installed on a participant's computer, the experimenter can change the behavior of the replacement new tab page by simply changing a parameter in a participant-specific file hosted on the experimenter's server. We used a fixed switching interval of 5 days, with all participants beginning with the default mode of new tab behavior, and with further switches assigned pseudo-randomly, ensuring that all participants received at least one block of 5 days' worth exposure to each of the three new behavior modes. Since we used a rolling recruitment process the 5-day blocks don't systematically interact with weekends for all participants, an important consideration in a behavior measurement experiment. This interaction is approximately random across subjects in our experiment.

2.2 Sample

Eighty volunteers from the general population (32 female, age 28.2 ± 6.1 years), recruited on a rolling basis, consented to participate in our experiment. 30% of them were university students, 40% were employed in the industry (excluding software industry), 20% were employed in the software industry and 10% were employed in the academia (lecturer/post-doc/professor). There were no web-developers in the sample, who used the browser for development. Participants were recruited through advertisements in newspapers, on social media websites and emails. Participants were not incentivized and were naive to the purpose of the study. We explained to them that we were interested in analyzing their web browsing history for a research experiment using a software extension they would install into their browser, that our extension would only store their data locally, and that they would have complete control over the curation process that would determine what part of their history they would share with the experimenters. The participants used various devices including laptops, desktops, tablet computers and mobile phones and no restriction on the type of device for a participant was imposed. A total of 10 participants reported to have used mobile phones / tablet computers while the remaining participants used either a laptop / desktop computer for installing the extension. At the end of two months of browsing for each participant, we asked them to curate the log files and share them with us, retaining the option to refuse. The two-month long browsing logs volunteered by these eighty participants formed our primary dataset.

At the time of submission of the logs, we asked the participants what fraction of the website logs had been curated by them. None of the eighty participants had curated more than 5% of the total content, with most participants not curating at all, which is reasonable and does not negatively affect our study. We had not disclosed to these participants that we intended to manipulate their new tab page's behavior remotely. To assess the extent to which they may have become aware of this possibility, at the time of collecting the logs we also asked them whether anything about the websites displayed in the new tab page 'bothered' them over the past two

months. They were asked to express their opinion on a 0-10 scale with 0 being the 'not aware' state and 10 the 'very aware' state. They were not specifically asked if they thought the type of websites shown were being altered remotely to avoid experimenter demand effects [6]. Subjects' responses are shown in Table 1.

Table 1: Participants' level of awareness of our experimental manipulation.

Opinion	0-4	5-7	8-10
Number of subjects	70	7	3

These survey results suggest that the majority of our participants remained naive to the purpose of the experiment, and continued to behave naturally in the face of our remote manipulations.

2.3 Data processing

The self-curated browsing logs obtained from all 80 participants in the experiment constituted our primary analysis material. These logs contained time-stamped instances of different website URLs visited by the participants. From these logs, we filtered out pop-ups and ads using a white-list, and used a 30-minute gap between page views to delimit browsing sessions. We collapsed web-pages visited on the same primary web domain into multiple visits to the same website, and likewise collapsed page refreshes (immediate repetitions of the same page in the log) as a single visit to the page. The few occurrences of dynamic web-pages in the browsing log were collapsed to that of the parent domain. We explicitly handled pages visited from search engines in the browser. For all occurrences of Google, Yahoo, Bing, Ask.com, Baidu, Wolframalpha and DuckDuckGo in the browsing log, we considered the page clicked immediately afterwards in the search results. If the search engine address was typed out in the new tab page, we treated the page clicked from the search results as if it was typed out in the address bar. Table 2 lists out a few anecdotal examples of the most visited and least visited websites across participants.

3 RESULTS

Since our hypothesis generally concerns the potential reduction of the diversity of websites that users visit as a consequence of the format of the new tab page display, our first coarse measurement of this property was the number of unique websites (defined via the primary web-domain in each URL) visited by our experiment participants during each mode of the new tab page's display properties.

Combining the default and most visited modes' data into one sample and for the blank and least visited modes' data into another, a two-sample T-test yields strong statistical significance $p < 10^{-3}$ and a very large effect size (Cohen's $d = 1.81$) as shown in Figure 2. Thus, there is clearly a large difference in user behavior across these experimental cohorts, in the predicted direction. Users visit a lot more unique websites on average when the new tab page is blank or shows infrequently visited websites, suggesting that the presence of recommendations by default is suppressing the repertoire of webpages they might naturally visit. For completeness, we show the results of bonferroni corrected pair-wise T-test among

Table 2: Representative examples of most visited (1st Quartile) and least visited (4th Quartile) websites of participants

Most Visited	Least Visited
Google	weebly
Facebook	Joomla
Youtube	icrc
Twitter	Netscape
Amazon	Scribd
Quora	Godaddy
Netflix	Nasa
Instagram	Yahoo
CNN	hotmail
Overleaf	Orkut

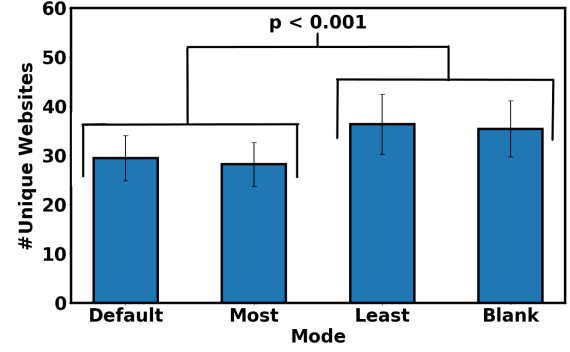
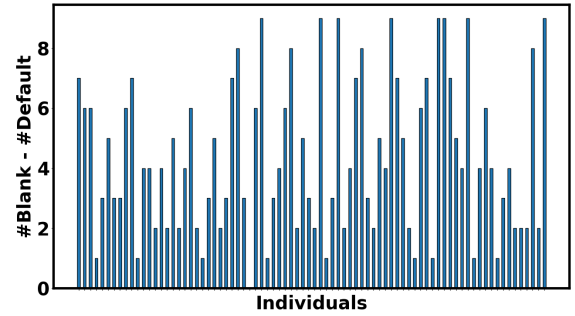
all the pairs of conditions in our study (Table 3). However, the

Table 3: Results of bonferroni corrected pairwise T-test for all the pairs of conditions (as shown in Fig. 3). Number of unique website counts for each of the four modes are averaged over 80 individuals.

Condition 1	Condition 2	p-value
Most visited	Least visited	0.000011
Most visited	Default UI	0.620102
Most visited	Blank UI	0.000008
Least visited	Default UI	0.000024
Least visited	Blank UI	0.522001
Default UI	Blank UI	0.000018

heterogeneity of web browsing behavior across individuals could conceivably inflate these statistics, in the sense that a few highly prolific web users in our participant pool could skew the cohort-level differences should they happen to conform to our hypothesis. Thus, it is important to also analyze the data at the individual-level. We plot the difference between behavior during blank UI mode and default mode for all participants in Figure 3. Note that not a single of the eighty participants visited more unique websites while the new tab page behaved under its default settings, vis-a-vis when the default recommendation display was switched off (Blank UI mode). A one-sample T test of these absolute differences with respect to zero sample mean was statistically significant at $p < 10^{-6}$. The average participant's unique site visit count increased by 15% (median improvement 12%) when the new tab page was set to display a blank page over their corresponding counts when the new tab page worked by default.

It is reasonable to conclude from this observation that changing the behavior of the new tab page does in fact alter the diversity of websites browsed. In particular, enabling recommendations in the new tab page, as per the default frequency scheme causes a reduction in the diversity of information sources (websites) accessed by the users. But this analysis does not substantiate our hypothesis

**Figure 2: Number of unique website count for each of the four modes averaged over all 80 individuals. Error bars represent ± 1 Standard Deviation (S.D.).****Figure 3: Absolute difference between the number of unique websites visited under the blank new tab page mode and under the default mode for all experiment participants.**

about the mechanism responsible for this reduction in diversity - an increase in users' proclivity to concentrate web browsing to familiar sources, caused by interaction with the new tab page interface.

We needed a more granular view of the data to see if this might actually be true. Recall that we store both the timestamped page visit, and the web event that brings the user to that page. We categorized the pages according to visit frequency and transition type to focus especially on page visits where we can analyze the nuances of users' visits to commonly visited and uncommonly visited websites and understand the implications of altering the new tab page display on the same.

This categorization is illustrated with a pseudo-flowchart in Figure 4, showing visually how we use the two decision points (frequency and transition type) to categorize web page visits into non-exclusive categories. To interpret the set-theoretic implications of the flowchart correctly, both *UI* and *pure* uncommon are subsets of uncommon, but *uncommon* and *common* are mutually exclusive. A brief description of these terms and their rationale follows.

- **Common.** Websites with high visit counts (removing the influence of interstitial URLs) are expected to correspond to content that users are in the habit of accessing and which correspond to commonly visited information sources. Since

the number of webpages an individual visits will inform what constitutes high for them, we categorize as *common* visits all page visits to webpages whose visit frequency occurs in the fourth quartile (Q_4) of visit frequencies sorted in ascending order per browsing history log. In a measure of the heavy skew in the distribution of webpage counts, this quartile accounts for about 70% of all site visits across all participants' data. A distinct sub-category "pure common" takes into account frequently visited webpages visited by typed transitions from the address bar (or search the web box) of an existing tab.

- **Uncommon.** As above, it is intuitive that when people are browsing websites that are not commonly visited by them, such browsing will be marked by visiting new sources of content, hence new webpages. We mark as *uncommon* visits all visits to pages that fall in the first quartile (Q_1) of visit frequencies sorted in ascending order per browsing history log¹. This quartile accounts for about 9% of all site visits across all participants' data.
- **Content / Event-driven.** Since we are interested in the interaction between the user's propensity to visit diverse information sources and the browser interface, we want to remove from our analysis all webpage visits that originate from an information source outside of these two. To this end, we mark all hyperlink based visits (marked in our logs as having the transition type 'link') irrespective of their frequency rank as content-driven, and exclude them from our analysis.
- **UI common.** The presence of clickable page icons obviously makes those webpages more accessible to users, over and above their mental propensity to visit these sites. Websites that users visit frequently, but use the UI to access, constitute an interesting sub-category of the *common* category. For these webpages, it is not as clear as for others that repeated visitation is indicative of conditioning due to passive recommendations that appear in the new tab page. The increased accessibility of the pages is clearly, but unquantifiably, also a factor. We refer to this category of pages, selecting all common pages arrived at either via UI clicks or by typing in the URL bar while the specific page icon was visible in the new tab page, as *UI common* pages.
- **UI uncommon.** When users visit rarely visited webpages via interaction with UI elements, we can infer that the UI interaction is driving them away from their typical sources of information. This makes the subset of *uncommon* pages that users arrive at from the new tab page (whether with transition 'link' or 'typed'), especially interesting. We call this sub-category *UI uncommon*.
- **Pure uncommon.** The final category we define is, expectedly, sparsely populated, representing about 3% of all web page visits in our dataset. We define this as the subset of

uncommon pages whose icons are not displayed on the new tab page, have a transition type of 'typed' and have been transitioned to directly from the new tab page. To be a member of this set, the user has to have entered the webpage URL (at least its first few letters, keeping in mind auto-complete capabilities of browsers) with no input from the browser UI. Visiting rarely visited sites without the (measured) influence of either other websites' content or the content of UI elements means that this subset of webpage visits reflect visits to uncommon sources of information driven purely by the interaction of users' internal preferences for URLs with the browser's new tab page. Privileging users' internal URL preferences in a semantic sense, we label this category *pure uncommon*. Within this category, we make two further distinctions: *manipulated* pure uncommon takes into account only typed transitions to rarely visited pages from the new tab page. *Un-manipulated* pure uncommon takes into account typed transitions to rarely visited pages from the address bar (or web search box) from an existing tab.

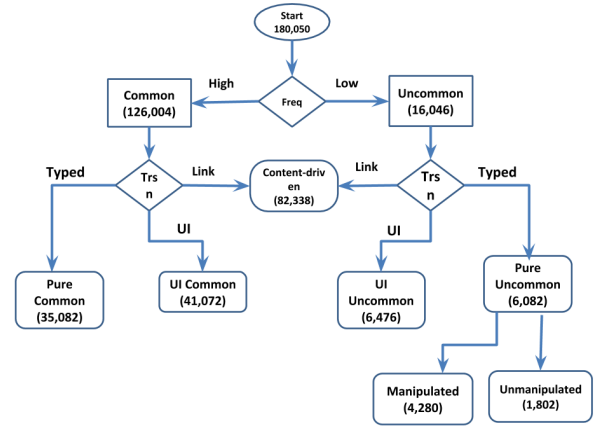


Figure 4: The analysis flowchart. The two decision points (diamond boxes) represent decisions by frequency (high = 4th quartile, low = 1st quartile) and transition type (link, typed, UI). Numbers inside boxes represent the number of webpage visits categorized under each category.

Using this categorization of browsing events, we sought to quantify the category-wise change in users' web browsing behavior across the three different UI manipulations. Because the different categories defined above have vastly different base rates, it is most sensible to measure the change across UI modes in terms of the category-wise percentage change,

$$100 \times \left(\frac{\text{New-category-occurrence-frequency}}{\text{Default-category-occurrence-frequency}} - 1 \right),$$

where the category occurrence frequency is mode-sensitive and is simply the number of page visits marked in a particular category for a user divided by their total page visits while their browser's new tab page's behavior was in a particular mode.

¹Not all visits to infrequently visited webpages imply a correlation with the passive recommendations in the new tab display. Sometimes users may have only need to go to a website, like a bank website very infrequently, on purpose. Thus, our definition overestimates the amount of 'diversity' as measured by this metric. Since we are ultimately interested, as we will see below, in the change in the amount of diversity of browsed web sources under controlled manipulation, this overestimation of the absolute quantity ends up not being crucial.

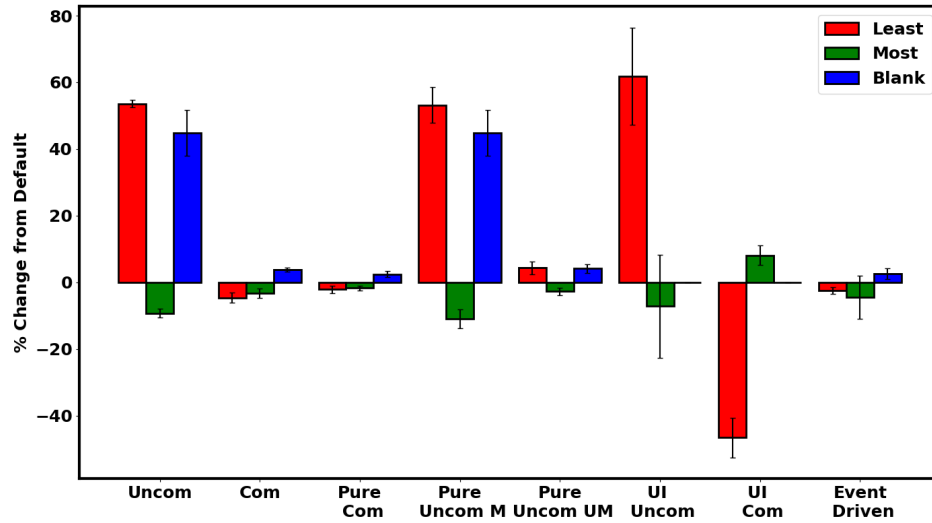


Figure 5: The change in each category’s occurrence fraction with respect to the same fraction seen in the default mode per subject, averaged across all subjects. Error bars represent ± 1 S.D. Com and uncom respectively denote the *common* and *uncommon* categories.

If the presence of web page recommendations on the new tab page affects browsing behavior by strongly concentrating visits to familiar information sources, we expected to see that our uncommon-linked categories would show positive percentage changes with respect to the default condition when we remove all page icons from the new tab display (the ‘empty UI’ condition) and when we present users with their least visited webpages as icons in the new tab display (the ‘least visited’ condition). As Figure 5 shows, this prediction is borne out clearly by our data. For both these conditions (least visited and empty UI), uncommon and pure uncommon categories showed large positive changes ($\sim 50\%$) with respect to the baseline condition (the default frequency-based UI web page icons display). These conclusions are statistically significant in Bonferroni-corrected pairwise T-tests ($p < 0.001$), with effect sizes $\{0.83, 0.89\}$ for uncommon and $\{0.87, 0.85\}$ for pure uncommon, for the least visited and blank UI display with respect to the default condition respectively.

Further, note also in Figure 5 that pages in the *manipulated* pure uncommon category accounted for almost all the change in the pure uncommon category as a whole; there are only marginal increases in the *non-manipulated* pure uncommon category. The difference between these two categories was only that the manipulated category page visits started with typing in URLs while the new tab page recommendations were visible, whereas the non-manipulated category page visits started from other pages, with new tab recommendations not visible. Since all other aspects of the users’ experience were constant across these two sub-categories, the stark difference in browsing behavior emerges only by virtue of the new tab page recommendations being visible or invisible. This performance dissociation allows us to confidently pinpoint the source of the difference in browsing behavior - the recommendations on the new tab page - with precision, and to infer its psychological

provenance also - interference in memory while retrieving website URLs to type in the URL bar by alternative displayed page icons.

Thus, we conclude that (a) the presence of new tab page recommendations causes people to visit fewer unique web pages than they do in the absence of new tab page recommendations, (b) the reduction in unique web page visits is most evident for websites users visit infrequently, and (c) a very large fraction of the reduction is caused by memory interference when people try to type in URLs on the new tab page.

4 QUANTIFYING THE CONSEQUENCES OF SOURCE CONCENTRATION

Our experiment shows that users’ web browsing behavior adapts to manipulations of the browser new tab display along lines predicted by theories of filter bubble and shallower processing of information on the web, and that this adaptation is *statistically* significant both across and within individuals. We have not, however, established that this adaptation is also *practically* significant. *Prima facie*, the changes in behavior are only apparent at the margins, not in the large aggregate of browsing patterns in our subjects, which comprises overwhelmingly of habitual access to highly visited websites. Viewed from this aggregate perspective, the deviations we have managed to introduce are of the order of about 0.5% of all page visits. Why should this miniscule change in behavior be consequential? This is the question we have sought to answer using the *in silico* experiments we report below.

Our primary hypothesis in this phase of the project was that minor changes in propensity to visit diverse information sources would affect the facility with which users can access crucial transition nodes in their information network that connect them to fresh sources of information that they do not commonly encounter during routine browsing. This hypothesis was inspired by related

observations about the role of social preference feedback in the polarization of internet communities - a phenomenon recently popularized by the term ‘filter bubble’ [24].

In classic filter bubble settings, the source of feedback and recommendations is external, but tuned to personal preferences. It is this personalization that, perversely, reduces the diversity of information sources that the user acquires their social media-mediated knowledge from [24]. In our case, unlike in conventional filter bubble settings, there are no external sources of information and/or feedback. The browser UI is simply passively reflecting the habitual behavior of the user, modulo some modelling of the inter-temporal preferences for various forms of content of users.

Nonetheless, we think the filtering effect of reinforcement of existing preferences works by the exact same mechanism in our case, suppressing the user’s propensity to sample content from diverse sources. Since the user’s response to their own previous history is much more observable than social influences seen in social media recommender systems, we can try to quantify the extent to which suppression effects of the size seen in our empirical data might influence the diversity of information sourcing for model agents.

To do this, we have modelled website browsing behavior for an individual agent as a random walk on a graph. Within the graph for each subject, each node represents a distinct website. With this abstract representation in place, the next step is to realize that the structure of the graph itself must reflect the Zipfian nature of webpage-website counts - users visits a small number of sites a large number of times, and a large number of sites a small number of times [14].

To capture these particular dynamics while maintaining analytic simplicity, we specialized the network structure we consider to be the family of all graphs with two pseudo-cliques wherein the cross-links between these pseudo-cliques are controlled by an ϵ parameter such that with a probability of ϵ , a link gets set up between two cross-pseudoclique nodes². When we simulate a random walk on any one such graph, the reasons for this construction become clear. Nodes within pseudo-cliques have several short paths connecting them, and so are likelier to be visited by a random walker known to be situated at another node within the clique. Nodes in different pseudo-cliques will have longer shortest paths on average, so depending on the distribution of length of random walks for the agent, visits across pseudo-cliques will be rarer. Thus, we induce differences in accessibility of different nodes in the graph by assuming its structure to take this particular shape.

Unlike more intricate agent-based models of browsing behavior, such as in [14], which try to capture both within and across website browsing behavior, the model we use above makes predictions only for the quantity we are interested in - the across website browsing behavior of the user. Dwell times and page depths within websites are not modelled in our approach. We offload as much of the preference information about browsing behavior as possible

onto the graph structure, resulting in the simplest possible observer model for the user - a random walk on the graph - that retrieves the Zipfian cross-site browsing behavior we are interested in simulating. The key difference between our simplistic model and the actual model of the web in terms of structure is that there are multiple pseudo-cliques in the web corresponding to different localizations of information [8]. The results we demonstrate for two pseudo-cliques can be shown for any pairs of pseudo-cliques without loss of generality, thus generalizing to more complicated web models as in [14].

The core of our succeeding analysis is identifying the slope of the relationship between the ϵ parameter, which controls the number of cross-links between the two pseudo-cliques, and the average first hitting time between nodes in different pseudo-cliques. Such an analysis does not require that the graphs contain pseudo-cliques of only equal size, or that they contain only two pseudo-cliques. We consider the simplest subset that instantiates the distance asymmetry we require. As a consequence of this distance asymmetry, we are able to map user behavior to network structure. Movements within a clique correspond to visitation of commonly accessed information sources while a leap from one clique to another corresponds to uncommon visits. Embedding the usually common, occasionally uncommon visitation behavior of users in this way in our graph model, we can say formally that there is on average uncommon visitation behavior ϵ of the time and visit to commonly accessed sources $1 - \epsilon$ of the time, assuming random initial node initialization in our specific graph structure. The same general principle is expected to hold for a much more general family of graphs - graphs with k pseudo-cliques, interconnected by cross-links generated via a stochastic process mediated by the parameter matrix $\epsilon_{ij}, i, j \in \{1, \dots, k\}$.

In this setup we are interested, fundamentally, in quantitatively characterizing the relationship between changes in ϵ and changes in the average first hitting time between cross-pseudoclique nodes. The first hitting time between two nodes is simply the expected number of moves a random walker originating at node i will take to first reach node j across multiple random walks originating at i . Since the UI manipulation affects ϵ by a large relative amount but a small absolute amount, we are interested in measuring the extent to which this small parametric change affects the visibility of cross-pseudoclique nodes in the graph, which corresponds in our browsing user model to the average first hitting time. Users will access websites that lie within their radius of experience, as measured by hitting time. First hitting times larger than the typical radius of experience - measured in terms of walk length - will correspond to websites typically inaccessible to the user.

Although theoretical properties of first hitting times on graphs have been investigated for random graphs [15, 31, 35], no direct results are specifically relevant for our analysis. Known theory suggests, on average, that adding extra nodes or edge to a graph increases the first hitting time of any two nodes in the graph but such analyses are true only on average across random graphs [2]. We, on the other hand, are interested in graphs with a very specific direction of variation in structure, changing from barbell style graphs for low values of ϵ into well-connected graphs as ϵ rises. A figure showing various graphs generated corresponding to different ϵ values is present in the supplementary material. The important

²A pseudo-clique is defined as the graph structure obtained by removing some edges from a clique. We generate pseudo-cliques by first assuming an exponential degree distribution within a pseudo-clique of nodes, then assigning edges to nodes randomly in decreasing order of degree, permitting ± 1 changes in the original degree assigned to each node to handle conflicts in edge assignment. In our experiments, we used pseudo-cliques of size $N=10$ with highest degree 4.

thing to note is the rapidity with which adding cross-links changes sample graphs drawn generatively using different values of ϵ from barbell style graphs (known to have worst case shortest path distances) to well-connected graphs, with much smaller shortest path distances. Even though the graphs with higher ϵ have more edges, we intuitively expect that the non-random manner in which the new edges are introduced by our graph generative model counteracts the mere fact that they are being added, and reduces the first hitting time.

The generative model for our graph has two free parameters, the size of the graph N and the cross-linking parameter ϵ . We fit this generative model to our data by fixing the size of the graph and the value of the ϵ parameter using the typical size of our users' unique website repertoire and their measured non-habitual behavior/habitual behavior ratios in our dataset respectively. The modal value of N (the number of unique sites visited by a user in a month) is 25. To approximate this, each of our simulated graphs used 20 nodes divided into two pseudo-cliques of equal sizes. We generated graphs using ϵ values varied between 0 and 0.2 in steps of 0.005.

Once an (N, ϵ) graph for a particular value of epsilon is ready, we empirically measured the hitting time of every cross-pseudoclique node pair by simulating random walks between all cross-pseudoclique node pairs. We calculated the average first hitting time for 100 such random walks across all cross-pseudoclique node pairs for each of 100 (N, ϵ) generated graphs for each ϵ value. Figure 6a shows the result of this simulation, plotting average first hitting times obtained from this experiment for different values of ϵ .

We found empirically that the range of ϵ , measured as the ratio of uncommonly visited to commonly visited pages for each subject, is 0.005–0.07 across our subjects. In this range, as illustrated in Figure 6a we observe an approximately exponential decline in hitting time with increase in ϵ . This is because, as anticipated, the effect of the 'cliqueish' structure of the graph dominates over increase in number of edges with ϵ . Gradually, as ϵ increases, the effect of increase in number of edges between pseudo-cliques starts dominating and hence the average hitting time asymptotes beyond $\epsilon \approx 0.1$, and will likely increase for still higher values. The key finding, also as anticipated, is that within the range of parametric values seen in our experiment, even small changes in ϵ cause large changes in the first hitting time. As we describe above, first hitting times represent website accessibility in our model, so the result obtained here implies that small changes in ϵ cause exponential declines in the accessibility of infrequently visited (cross-pseudoclique) websites for users.

Finally, the relationship between a finite random walk and the probability that it will pass through any given node is straightforward to estimate. It is simply the probability that a random walk as long as N might occur in the empirical distribution of first hitting times for that specific node. Hence, the p-value of a one-sided probability hypothesis test will give us exactly the probabilistic quantity of most interest to us - the probability with which a user with non-habitual browsing propensity ϵ will access a diverse information source in any given browsing session of any particular estimated length.

We identified individual sessions in our participants' browsing logs using a 30 minute interval to determine the boundaries of individual sessions (as mentioned in Section 2.3), and counting total

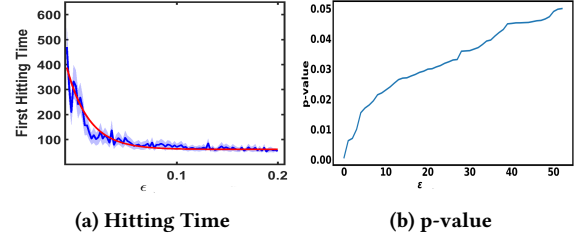


Figure 6: (a) Plot of average and 95% confidence bound of first hitting time for all cross-pseudoclique nodes empirically measured across 100 simulated random walks in each of 100 graphs generated for each value of ϵ . The best fit exponential for the data is also shown. (b) Variation of p-value of the average session length with ϵ . p-value is averaged across multiple cross-pseudoclique node-pairs and graphs.

page visits as one step in the random walk. Across all participants, the average session length came out to be around 20 for our dataset. Next, we measured the kurtosis of the histogram of hitting times for each node pair in multiple random walks on a particular graph. We found that the kurtosis values themselves were well-represented by a normal distribution with mean 3.1 and SD 0.2. Hence, the first hitting time distribution for each node pair could be considered approximately Gaussian.

So, we ran one-sided z-tests testing whether a session length of $k = 20$ or lower might occur naturally in the hitting time distribution of cross-clique node pairs. We ran this calculation for 100 randomly sampled cross-pseudoclique node pairs across 100 graphs generated for each value of ϵ and plotted the p-values obtained against ϵ . This relationship is visualized in Figure 6b, and is approximately linear. This observation quantifies concretely the importance of large relative changes in the propensity to visit diverse information sources, even if this propensity constitutes a small relative share of overall browsing behavior. While the change in behavior is on the margin in terms of aggregate browsing behavior, it has non-marginal consequences on which information sources users will be able to access.

5 DISCUSSION

Since the beginning of internet browsing, designers have always tried to design browsers in ways that reduce the user's typing effort. In early browsers, this was done using history auto-complete suggestions. More recently, both OS and browser designers have sought to reduce typing effort still further by adding apps to home screens and frequent (and recent) web pages to 'new tab' pages. The inarguable logic of such design is that, because website visit counts are approximately power law distributed, being able to simply click on the most visited pages optimizes the number of typing responses needed over the course of users' use of the browser.

But website visits are a means to an end, not an end in themselves. Ultimately, they are expressions of our information preferences [25]. This paper makes the case that, by showing people the sites they visit most frequently over and over again in new tab displays, current practice in browser and OS UI design traps

people into a solipsistic feedback loop, reinforcing their strongest preferences to the detriment of weaker ones that potentially offer more scope for diversity of experience and learning [29, 33]. We quantified the size of this feedback effect at its source - natural browsing behavior - by collecting browsing history from subjects who consented to having the behavior of their new tab page manipulated remotely via a plug-in they installed in their browsers. With *in silico* experiments, we further demonstrated the impact of the change in browsing behavior on peoples' proclivity for acquiring information from diverse sources.

5.1 Related work

Our work follows a rich vein of empirical research in human-computer interaction that is increasingly discovering subtle but powerful ways in which website interaction, access and even simple UI decisions can influence human decisions. For example, posting to Facebook has been shown to increase users' activity on the website immediately before and immediately after the posting, on the timescale of days [12]. More substantively, time of day of Twitter activity has been shown, strikingly, to correlate strongly with the probability of users being clinically depressed [7]. Even more strikingly, Epstein and colleagues have shown that presenting potential voters with search results for potential candidates in an actual national election in a UI with artificially manipulated search result rankings can shift the voting pool's vote fraction by up to 2% in the experimenter's chosen direction, a shift large enough to sway a reasonably close election [9]. There have also been recent works studying memory recall in simple tasks like clicking of pictures through different capture modalities [23]. We identify a large and consequential psychological effect: conventional UI design choices suppress peoples' propensity to access diverse information sources while accessing the web. This finding is not as specific as the ones documented previously in the literature, but also as a consequence, is likely affecting a lot more people at any given point in time by virtue of its generality.

Our results also relate to research efforts ongoing to characterize and surmount the recommender systems (RS) *filter bubble* [4, 18, 28]. There is mixed evidence regarding the effects of recommender systems on users' consumption patterns. Whereas work focused on identifying changes in content *supply* as a function of personalization tools has found no effects [13], other work studying *consumption* patterns have reported both negative [4, 10] and positive effects on consumption diversity [22]. Our work finds unambiguous and large negative effects on information diversity driven by the use of recommendations. By our simplest measure of diversity - count of unique websites visited - users operating browsers with blank new tab pages visited on an average 15% more unique websites over a two month period.

One possible reason for previously reported null or positive results in the literature might be that the *active* use of recommender systems like Movielens involves users actively looking for information they don't already possess, an explicit cognitive task where they are aware that they must evaluate the incoming information quasi-critically [30]. In contrast, constantly stopping by the new tab page during transitions between websites is a much subtler phenomenon; the recommendations here are *passive*, in the sense that the ostensible function of these bookmarks is to facilitate access to

frequently visited pages, not to shift preferences. People can reason through, or actively ignore, redundant information while they are actively deciding what to do. It is harder for them to guard against the subtle impact of memory inhibition through visual presentation of web page icons time after time in the normal flow of their interaction with the browser [21].

While it might seem superficially strange that such a small UI design component could potentially have such a large impact on browsing behavior, behavioral economics offers striking examples of similar subconscious nudges significantly affecting behavior for better and for worse [16]. Similarly, Thaler and Sunstein offer striking examples of how small and often insignificant things can influence (nudge) behavior [17]. As McLuhan foresaw, but many modern empirical analyses ignore, the peril in informative computer interfaces is not that they may feed us bad information against our will, it is that the *form* in which they present information to use will subtly change our expectations of what we want to look for.

The classic ideal observer in web browsing - Pirolli & Card's *information forager* - models humans as observers optimizing information acquisition under information-theoretic constraints [25–27]. Our demonstration of large priming effects caused by as innocuous a source as the new tab page interface could help further nuance such models by incorporating the effect of psychological biases in the same way the influential cognitive biases literature has influenced the design of microeconomic models [34].

Our description of the delicate interaction between mind and machine in this most humdrum of tasks - web browsing - should also stimulate further exploration of the negative externalities of personalization, recommender systems and other methods of preference influence common in HCI applications. At present, such influence is sometimes justified by discriminating between malevolent and benevolent deception in UI design, with the basic difference being that a benevolent deception is meant to benefit the user [1]. Our study presents a clear example of the inadequacy of such ethical definitions - in our case, the developer, in principle, gains nothing by presenting frequently viewed web pages as clickable icons; the UI is designed the way it is to minimize the user's need to type. Nonetheless, this benevolent deception, conflating accessibility with preference via the subconscious impact of presentation on memory, ends up affecting user behavior substantially, as our results show.

5.2 Limitations

A natural limitation of this work is the size and representativeness of the population sample we used to establish the basic fact that presentation of commonly used websites' page icons on the new tab display interferes with users' proclivity to visit diverse information locales and concentrates web browsing to familiar sources. Nonetheless, the empirical results we describe in this paper are statistically significant by all conventional standards of measurements, and our sample is quite diverse with respect to gender, education, and occupational status. While one can never have too many participants in a study making claims about internet browsing behavior, the large effect sizes we find, consistently seen both across and within participants, suggest that our study was adequately powered to discern the specific hypothesis it investigated.

It is also possible to question how well conclusions from this study, with a mean sample age of 28, may generalize to the experience of older adults. At the same time, it must be remembered that while our sample is not representative of the general population, it is certainly more representative of *heavy* internet users [3, 32], who are the population most likely to be affected by the UI design decisions critiqued in this paper. We also note that the average age of our population is very close to the median age of the country this study was conducted in, suggesting that there is no significant age-bias in our recruitment.

5.3 Conclusion

In this paper, we contribute two basic results. One, we show empirically that subtle UI manipulations of the new tab page in common browsers can create large relative reductions in users' propensity to explore sites outside the repertoire of sites they commonly visit. Two, we show using a simulation study informed by empirically measured parameters that large relative changes in web source concentration manifest linearly in large absolute changes in the visibility of diverse information sources in a user's browsing experience. Together, these results present striking evidence that the current design practice of showing frequently visited webpages on new tab pages is suppressing the expression of web surfers' tendencies to visit diverse information sources on the web.

Considering how ubiquitous computers and the internet have become, this effect would span even mobile devices. In fact, since small screen sizes in hand held mobile phones or tablets are likely to bias users towards clicking the new tab page icons over typing in the address bar, the concentration of web browsing to familiar sources is likely to be exacerbated, as an admittedly small *post hoc* cohort analysis of our data shows. Further research is needed to quantify the extent of this difference.

Finally, we note that all common browsers offer users the ability to use a blank page as their new tab page, making our conclusions immediately actionable. Based on our results, people should decide for themselves whether the convenience of clicking through to commonly visited web pages is worth the potential curtailment of their curiosity.

ACKNOWLEDGEMENT

We thank DST, Government of India for their funding in support of the project through grant DST/CSRI/2017/334 (G). We also thank all the anonymous reviewers for their valuable feedback.

REFERENCES

- [1] Eytan Adar, Desney S Tan, and Jaime Teevan. 2013. Benevolent deception in human computer interaction. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1863–1872.
- [2] David Aldous. 1989. An introduction to covering problems for random walks on graphs. *Journal of Theoretical Probability* 2, 1 (1989), 87–89.
- [3] Henry Assael. 2005. A demographic and psychographic profile of heavy internet users and users by type of internet usage. *Journal of advertising research* 45, 1 (2005), 93–123.
- [4] Eytan Bakshy, Solomon Messing, and Lada A Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132.
- [5] Nicholas Carr. 2010. *The shallows: How the internet is changing the way we think, read and remember*. Atlantic Books Ltd.
- [6] Gary Charness, Uri Gneezy, and Michael A Kuhn. 2012. Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization* 81, 1 (2012), 1–8.
- [7] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting Depression via Social Media. *ICWSM* 13 (2013), 1–10.
- [8] Zoltan Dezső, Eivind Almaas, András Lukács, Balázs Rácz, István Szakadát, and A-L Barabási. 2006. Dynamics of information access on the web. *Physical Review E* 73, 6 (2006), 066132.
- [9] Robert Epstein and Ronald E Robertson. 2015. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences* 112, 33 (2015), E4512–E4521.
- [10] Seth Flaxman, Sharad Goel, and Justin M Rao. 2016. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly* 80, S1 (2016), 298–320.
- [11] Mozilla Foundation. 2014. Frecency algorithm. Blog. Retrieved September 19, 2017 from https://developer.mozilla.org/en-US/docs/Mozilla/Tech/Places/Frecency_algorithm.
- [12] Nir Grinberg, P Alex Dow, Lada A Adamic, and Mor Naaman. 2016. Changes in engagement before and after posting to facebook. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 564–574.
- [13] Mario Haim, Andreas Graefe, and Hans-Bernd Brosius. 2018. Burst of the filter bubble? Effects of personalization on the diversity of Google News. *Digital Journalism* 6, 3 (2018), 330–343.
- [14] Bernardo A Huberman, Peter LT Pirolli, James E Pitkow, and Rajan M Lukose. 1998. Strong regularities in world wide web surfing. *Science* 280, 5360 (1998), 95–97.
- [15] Satoshi Ikeda, Izumi Kubo, and Masafumi Yamashita. 2009. The hitting and cover times of random walks on finite graphs using local degree information. *Theoretical Computer Science* 410, 1 (2009), 94–100.
- [16] Thomas C Leonard. 2008. Richard H. Thaler, Cass R. Sunstein, Nudge: Improving decisions about health, wealth, and happiness.
- [17] Thomas C Leonard. 2008. Richard H. Thaler, Cass R. Sunstein, Nudge: Improving decisions about health, wealth, and happiness.
- [18] Q Vera Liao and Wai-Tat Fu. 2013. Beyond the filter bubble: interactive effects of perceived threat and topic involvement on selective exposure to information. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2359–2368.
- [19] Marshall McLuhan, MARSHALL AUTOR MCLUHAN, and Lewis H Lapham. 1994. *Understanding media: The extensions of man*. MIT press.
- [20] Caroline Myrberg and Ninna Wiberg. 2015. Screen vs. paper: what is the difference for reading and learning? *Insights* 28, 2 (2015).
- [21] Ben R Newell and David R Shanks. 2014. Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences* 37, 1 (2014), 1–19.
- [22] Tien T Nguyen, Pik-Mai Hui, F Maxwell Harper, Loren Terveen, and Joseph A Konstan. 2014. Exploring the filter bubble: the effect of using recommender systems on content diversity. In *Proceedings of the 23rd international conference on World wide web*. ACM, 677–686.
- [23] Evangelos Niforatos, Caterina Cinel, Cathleen Cortis Mack, Marc Langheinrich, and Geoff Ward. 2017. Can Less be More?: Contrasting Limited, Unlimited, and Automatic Picture Capture for Augmenting Memory Recall. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 2 (2017), 21.
- [24] Eli Pariser. 2011. *The filter bubble: What the Internet is hiding from you*. Penguin UK.
- [25] Peter Pirolli and Stuart Card. 1995. Information foraging in information access environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 51–58.
- [26] Peter Pirolli and Stuart Card. 1999. Information foraging. *Psychological review* 106, 4 (1999), 643.
- [27] Peter Pirolli and Wai-Tat Fu. 2003. SNIF-ACT: A model of information foraging on the World Wide Web. *User modeling 2003* (2003), 146–146.
- [28] Paul Resnick, R Kelly Garrett, Travis Kriplean, Sean A Munson, and Natalie Jomini Stroud. 2013. Bursting your (filter) bubble: strategies for promoting diverse exposure. In *Proceedings of the 2013 conference on Computer supported cooperative work companion*. ACM, 95–100.
- [29] Jürgen Schmidhuber. 2010. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development* 2, 3 (2010), 230–247.
- [30] Rashmi Sinha and Kirsten Swearingen. 2002. The role of transparency in recommender systems. In *CHI'02 extended abstracts on Human factors in computing systems*. ACM, 830–831.
- [31] Vishal Sood, Sidney Redner, and Dani Ben-Avraham. 2004. First-passage properties of the Erdős–Rényi random graph. *Journal of Physics A: Mathematical and General* 38, 1 (2004), 109.
- [32] Lori C Soule, L Wayne Shell, and Betty A Kleen. 2003. Exploring Internet addiction: Demographic characteristics and stereotypes of heavy Internet users. *Journal of Computer Information Systems* 44, 1 (2003), 64–73.
- [33] Nisheeth Srivastava, Komal Kapoor, and Paul R Schrater. 2011. A cognitive basis for theories of intrinsic motivation. In *Development and Learning (ICDL), 2011 IEEE International Conference on*, Vol. 2. IEEE, 1–6.
- [34] Amos Tversky and Daniel Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *science* 185, 4157 (1974), 1124–1131.
- [35] Kôhei Uchiyama et al. 2011. The first hitting time of a single point for random walks. *Electronic Journal of Probability* 16 (2011), 1960–2000.