DESIGNING TECHNOLOGIES FOR EXPOSURE TO DIVERSE
OPINIONS

BY

QINGZI LIAO

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Computer Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

      Associate Professor Wai-Tat Fu, Chair
      Professor Alex Kirlik
      Professor Bruce Schatz
      Dr. Peter Pirolli, PARC

ABSTRACT

Exposure to diverse opinions can help individuals develop accurate beliefs, make better decisions, and become more understanding and tolerant persons. It is also necessary for the governance of a stable and democratic society. However, the exposure is often limited by people's natural tendency towards *selective exposure*—preferential seeking of confirmatory over challenging information. This has prompted many scholars to warn about the rise of "echo chambers" and "filter bubbles" online, with individuals' easy control over information exposure enabled by digital technologies. Such concern has motivated human-computer interaction (HCI) researchers to study a class of *diversity-enhancing technologies*—information and social technologies that host diverse viewpoints and take increasing users' exposure to information that challenges their existing beliefs as a design goal.

In this dissertation, I seek to answer the following question: *What kind of design features can nudge users to be exposed to more attitude-challenging information?* To complement the current technical-HCI approaches, I focus on bridging social science theories on selective exposure and design guidelines for diversity-enhancing technologies. Specifically, the central objective of this dissertation is to understand key factors that moderate individuals' propensity to engage in selective exposure in interacting with information and social technologies and apply the knowledge in four aspects of diversity-enhancing technology design: 1) design by enabling moderators that reduce, and eliminating ones that increase, selective exposure; 2) design for personalization by identifying user groups and use contexts that have varied selective exposure tendencies; 3) design for personalization by tailoring diversity-enhancing designs based on the underlying individual differences; and 4) design beyond individuals by considering the opinion group differences in selective exposure tendencies and the implication for user behaviors and social network

structure.

This dissertation provides empirical evidence that user behaviors in seeking attitude-relevant information are subject to the influence of various individual and contextual factors and recommends a more personalized approach that carefully controls and leverages these factors to nudge users into more desirable information consumption. It contributes several new lessons for designing technologies that present diverse viewpoints, including a theory-driven guideline for personalizing diversity-enhancing designs, insights on the selective exposure bias in consumer health information seeking, and an exploration of group selective exposure and its implication for social technology design. Perhaps most importantly, the dissertation pinpoints several directions in which selective exposure theories can be applied to the design of diversity-enhancing technologies, which opens up opportunities for developing a unified knowledge framework to push this research field forward.

Dedicated to my beloved parents,
who teach me to appreciate sciences and humanities.

ACKNOWLEDGMENTS

I owe this dissertation to many people who have provided feedback and advice on my research, and those who have kindly and patiently supported me over the years.

First and foremost, I am forever grateful to my advisor Professor Wai-Tat Fu. Thank you for taking me in the lab seven years ago, and guiding me from taking baby steps in research to who I am today. I am always learning from your brilliant mind and positive attitudes.

To my committee members. Professor Alex Kirlik was one of the first professors I met when I started studying in the Human Factors program for my Master degree. You have been a constant source of support and kept pushing me forward in my research by providing valuable insights and constructive feedback. Professor Bruce Schatz has always been an incredibly helpful mentor, who has encouraged me to pursue challenging research ideas, and offered countless advice and support for my research career. Dr. Peter Pirolli's work and ideas have always been a source of inspiration. It was truly an honor to have worked with you at PARC, and receive advice from you that tremendously helped me complete this dissertation.

I would also like to thank my other mentors, collaborators, and those who have aided me and shaped who I am as a researcher. To this end, I would like to thank Dr. Victoria Bellotti, Dr. Markus Strohmaier, Dr. Matthew Davis, Dr. Michael Muller, Dr. Michael Youngblood, Dr. Daniel Morrow, Professor Karrie Karahalios, Professor Brian Bailey, Professor Chengxiang Zhai, Dr. Claudia Wagner, Dr. Rajan Vaish, Dr. Yunliang Jiang, Dr. Jessie Chin, and Jason Cho.

I am also grateful to the constant support from my great colleagues over the years. To my current and former labmates, Dey Sanorita, John Lee, Eric

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

Our society has begun embracing diversity as a political and social norm. Exposure to dissimilar viewpoints is indispensable in developing accurate beliefs and making informed decisions. Without being challenged by alternative views, one may hold on to the wrong choice with illusory confidence—a problem that has been ubiquitously reported in decision-making in various domains, ranging from criminal investigation [2] and scientific research [3] to health care [4, 5], financial investment [6], education [7],and many other aspects of our everyday choices [8]. As John Stuart Mill famously put it, "(when people do not have sufficient exposure to different opinions), *if the opinion is right, they are deprived of the opportunity of exchanging error for truth. If wrong, they lose, what is almost as great a benefit, the clearer perception and livelier impression of truth, produced by its collision with error.*"

Cross-cutting exposure to diverse views is also considered a central element in public life for sustaining a democratic and stable society. Advocates of deliberative democracy deem the premise of democracy to be a "marketplace of ideas" — "*reasonable, carefully constructed, and morally justifiable arguments to one another in a context of mutual respects*" among people of diverse viewpoints [9, 10, 11]. Others emphasize the benefits of personal interactions with people of diverse backgrounds for developing empathy for ideological, social and cultural differences, eliminating undesirable stereotypes, and accepting the legitimacy of various political outcomes [12]. With the growing divide along the ideological spectrum, such as the division of American citizens into the ever more distant "blue" and "red" camps, communication across lines of difference has been increasingly sought after as the safeguard for ideological polarization and dangerous radicalization [13].

By providing vast enhances in information access as well as weakening the geographic and social boundaries, the Internet has, since its inception,

been expected to transform the "marketplace of ideas" into the virtual space, where opportunities for opinion expression, idea exchange, and public deliberation abound. In particular, many have been cheering for the promise offered by social media for Habermas' public sphere [14].

The availability of diverse information in the environment, however, does not guarantee that people's information consumption will be equally diverse. According to *selective exposure theory*, individuals are driven to seek information that supports their views and turn away from disagreeable information to avoid cognitive dissonance—the psychological discomfort arising from having to reconsider one's beliefs or decisions [15]. The implication of selective exposure is that, when individuals have the choice, they may live in an "echo chamber" where they are only repeatedly exposed to information that reinforces their existing views.

In the Internet era, with the declines of broadcasting media channels serving as gatekeepers of information for diverse audiences, individuals have gained unprecedented control over their own information exposure by easily selecting which website to visit and whom to follow on social media, and by using various information filters such as ranking algorithms, personalized user settings, self-curated information aggregators, etc. These "narrowcasting" channels may profoundly change people's information consumption, prompted many scholars to warn that the Internet may exacerbate selective exposure and lead to increasing social fragmentation [16, 17, 18]. Among these scholars, Cass Sunstein made the most compelling criticism in his book *Republic.com* [16], predicting that the Internet will inevitably make public discourse more fragmented. In the same vein, Eli Pariser, in his best-selling book *The Filter Bubbles: What the Internet Is Hiding from You*[19], vividly depicted living in one's own opinion and cultural "bubbles" as an outcome of conscious or unconscious use of popular personalization technologies that narrowly define personal preference as agreement, making one close-minded and potentially vulnerable to propaganda and manipulation.

1.1 Diversity-Enhancing Technology Design: Existing HCI Approaches to Mitigate Selective Exposure

The alarming implications of digital media for selective exposure have motivated a small but growing group of human-computer interaction (HCI) researchers to study how to design technologies that can diversify users' information diet, in order to support learning, decision-making, personal reflection, and public deliberation. In this dissertation, I use the term "**diversity-enhancing technology**" to refer to a class of tools that host diverse viewpoints, and consider enhancing the exposure to challenging information as a design goal.

The scope of diversity-enhancing technology is broad, including but not limited to the following: 1) Deliberation platforms that support collective and democratic decision-making, where a group of users presents and discusses diverse viewpoints with the goal of reaching consensus, as exemplified by the LiquidFeedback platform[1]; 2) Deliberation platforms that support individual learning and reflection on complex issues. Examples include Consider.It [20] and OpinionSpace[21], both focusing on helping individuals navigate through a diverse collection of comments from other people; 3) Decision-support systems (DSS) that focus on remedying biased seeking of confirmatory information, which often adopts various de-biasing techniques. A typical example is a clinical decision-support system that encourages physicians to consider and compare alternative options (e.g., [22]); 4) News and other information aggregators (e.g., social media feed) designed with the goal of enabling individuals to continuously obtain information from diverse information sources; and 5) Social opinion websites (e.g, Yelp), and, more broadly, online communities, where one may seek diverse opinions for current issues, products, services, and many other topics.

The existing HCI effort for diversity-enhancing technologies has been primarily technical, focusing on developing new platforms or novel interaction techniques that support navigation of, and deliberation on, diverse views. For example, developing political deliberation platforms is one such area that is fuitful. A number of visualization techniques have been explored to visually present the arguments, popularity, and interactions between different stances

---
[1]http://liquidfeedback.org/

3

[21, 23, 24, 25]. Many platforms, e.g., NewsCube[26], OpinionNetIt[27], and Poli[28], focus on building structured interfaces (e.g., faceted interface) to present diverse opinions, while Consider.It[20], Reflect[29], OpinionSpace[21] and some other tools [30] proposed novel techniques that encourage active user actions to explore and collect information from diverse viewpoints.

Because of the novel nature of these tools, they are often evaluated qualitatively or using simple self-reported measurements. Although favorable opinions of the systems were reported, direct assessments on reducing selective exposure *per se* have rarely been conducted. While such practice may arise from seeing diversity exposure as a means to the ends of deliberation and system success, as a result, we still lack a fundamental understanding of *what kind of design features can effectively increase the diversity of information people consume.* This means that, for one, there lacks unified knowledge to account for the varied empirical results and generate new ideas for future system designs. Moreover, we may not be able to transfer the successful design examples to other systems, especially systems that are not specialized in supporting deliberation. For example, given that many criticize popular social media like Facebook as magnifying users' selective exposure tendency, how should we introduce new, or leverage existing, design features to mitigate the problem? There is the question of transferability to other domains as well. For example, while several studies reported confirmatory information seeking as a form of bias problem in consumers' online health information seeking [31, 32, 33], how do we know what kind of diversity-enhancing design would be effective and well-received in this new context?

This dissertation is not the only one that expresses and responds to such concern. Very recently, a panel discussion was held at the 2013 ACM CSCW conference [34], calling for identifying more general strategies that can promote exposure to diverse opinions. These strategies were referred to as "nudges," hailing to the concept originally developed by behavioral economists Richard Thaler and Cass Sunstein [35]. A nudge is an intervention built in the environment that can change people's behavior in desirable directions without limiting their freedom of choices, often by enlisting knowledge from behavioral and social sciences. For designs that aim to reduce selective exposure, a nudge can range from a subtle design feature that changes users' information-seeking paths to novel interaction techniques that generate new

behavioral patterns.

Thaler and Sunstein called the kind of behavioral and social science research that studies interventions for behavior change as efforts towards building a "toolbox of nudges." In the same vein, this dissertation represents an effort to inform a *"toolbox of nudges for designing diversity-enhancing technologies."* This means a shift from the technical focus to *understanding people* and how their behaviors can be changed in the direction of increased exposure to diverse views. A key opportunity, as argued by Garrett and Resnick (2011), lies in understanding what factors would enhance or reduce individuals' selective exposure tendency and, with that, to enable *personalization design*—not the kind of personalization narrowly prioritizing user agreement and creating "filter bubbles", but leveraging numerous personal and contextual factors to increase individual tendency to seek different opinions in specific settings.

1.2   My Approach: Designing Diversity-Enhancing Technology with Moderators of Selective Exposure

The motivation for a personalized design approach to tackling selective exposure challenges has both empirical and theoretical foundations. Several empirical studies examining people's online information behaviors reported that preferences for diversity vary across individuals and information-seeking contexts. Although a general preference for confirmatory information exists, evidence shows that some users do actively look for information across the ideological spectrum [36, 37], and in some contexts, especially non-political ones [38], individuals may value the utility of alternative views for improving learning and decision outcomes [39, 40].

While most of these observations are merely empirical without concluding the reasons behind the individual and contextual differences, we must be aware that studying moderators of selective exposure—what variables increase or decrease selective exposure tendency—has been a longstanding focus of psychological research on selective exposure theory. This can be traced back to the 1960s when the early critics of selective exposure theory ques-

tioned the existence of an underlying psychological preference for supportive information [41], and pointed to the mounting evidence failing to observe selective exposure in experimental research. Since the late 1980s, researchers in this field started to converge on studying the moderators of selective exposure to account for the contradicting evidence and, more importantly, to deepen our understanding on attitude relevant information seeking.

One of the most prominent efforts to date is Hart et al.'s work [42]. By conducting a meta-review of 91 studies on selective exposure, they presented a general framework using two opposing motivational constructs—defense motivation and accuracy motivation—to encompass the numerous moderators identified in previous research. Specifically, positive moderators of (increasing) selective exposure, such as self-threat [43, 44] and commitment to one's position [45, 46], are recognized to incur defense motivation—desire to defend one's position, leading to more pronounced selective exposure. Negative moderators of (decreasing) selective exposure, such as outcome importance [47] and information utility for learning goals [48], are considered to foster accuracy motivation—desire to form accurate understanding of the topic, which promotes tendencies to seek information in an unbiased fashion.

Although the rich literature on moderators of selective exposure could be a treasure that has not yet been unearthed for designing diversity-enhancing technologies, there are certainly many questions that need to be answered for the two fields to be bridged. First, naturally we need to inquire the *transferability* of knowledge from traditional lab studies to technological environments. This is not only because the former tended to use a paper-based information environment with limited information choices but also because they largely followed a post-decision information-search paradigm to study selective exposure, which may not be the primary use case for interacting with diversity-enhancing technologies. Instead, our regular, often casual, use of information and social technologies that present diverse views tends to involve more browsing than goal-oriented searching, activities, and we are often placed in an information-rich environment where access to both similar and dissimilar views is just a few clicks away.

A second and perhaps more interesting question is on the *application* of the knowledge about moderators of selective exposure from behavioral science research to the designs of diversity-enhancing technologies. We certainly

should not stop at knowing which users may be more open or resistant to dissimilar opinions, but we ought to consider how to apply knowledge about the moderators to active engineering of the system interfaces. In this dissertation, I seek to understand the key moderators of selective exposure in interacting with information and social technologies and argue that they can be applied to four goals that are the keys to the success of diversity-enhancing technologies. In Table 1.1, I list the following goals and how they are covered in the chapters of this dissertation, each of which presents one study I conducted.

- **Goal 1:** *Design nudges with moderators* to **enable ones that reduce selective exposure, and eliminate ones that exacerbate selective exposure**, in order to nudge users towards exposure to diverse viewpoints. For example, based on the accuracy-defense motivation framework, many design ideas can be generated to accentuate the benefits of, or enable, the accurate understanding of the topic or to alleviate users' incitement for defense. While this goal is rather straightforward, I emphasize the focus on changing interface features that can induce the known moderators of selective exposure. In the dissertation, to exemplify this approach, using an information aggregator for controversial topics, I showed that design features as simple as a threatening image can unintentionally trigger users' defense motivation, and thus should be cautiously avoided (Study 1 in Chapter 3). I also explored several design principles for enhancing diversity exposure based on moderators suggested by selective exposure theory (Studies 2-4 in Chapters 4-6).

- **Goal 2:** *Personalization design with moderators*, which may involve two steps. The first is to **identify user groups and use contexts with varied selective exposure tendencies** that need to be personalized for. While personalization tasks can be complex and involve numerous factors, the accuracy-defense motivation framework provides a high-level construct to predict user groups that are more open to a diverse collection of information and ones that are averse to challenging information and thus need to be more closely targeted to mitigate selective exposure. For example, based on the accuracy

motivation construct, we may expect user groups that are genuinely interested in the topic to be less prone to selective exposure. Based on the defense motivation, we may predict user groups that are frequently challenged by others to develop a stronger tendency to seek confirmation. This applies to adapting systems for different use contexts as well by predicting whether a user may become more or less likely to engage in selective exposure for different topics or when encountering certain events. Ultimately, a comprehensive list of moderators can enable accurate prediction of users' diversity preference in a dynamic fashion, which may serve as the foundation for a more intelligent and adaptive "nudging" approach. Given that identifying a complete list of moderators is beyond the scope of any dissertation, my research focused on exploring the feasibility of this theory-driven approach. Specifically, I validated that topical involvement, as a known contributor of accuracy motivation, led to more balanced information seeking when interacting with an information aggregator (Study 1 in Chapter 3). By studying selective exposure with real user-data from social media discussions on a controversial topic (Study 5 in Chapter 7) and by studying it in a different, less studied domain—consumer health information seeking (Study 4 in Chapter 6)—I showed that heightened threat, a factor repeatedly reported to be associated with defense motivation [49, 50, 51], led to more pronounced selective exposure in interacting with technologies.

- **Goal 3:** A second step in ***personalization design with moderators*** is to **tailor nudging designs for different user groups and use contexts** to achieve the optimal diversity-enhancing results. While Goal 2 is to answer the question of "for whom to personalize," Goal 3 is to answer the "how to personalize" question. It is a necessary question as experiment showed that not only does there exist a division of *diversity-seeking* and *challenge-averse* users, but they also tend to respond differently to diversity-enhancing designs [36]. Behavioral science research on moderators of selective exposure can inform this goal by revealing the underlying factor that contributes to the observed user differences. For example, by identifying difference in accuracy motivation as the underlying factor that mediates users' information behaviors, we may seek design solutions based on the needs and preferences

resulting from varied levels of accuracy motivation. Specifically, in this dissertation, by associating accuracy motivation with the rich literature on dual-process persuasion theories [1] and their applications in tailoring behavior-change interventions based on motivational levels, I proposed a theory-based design guideline to tailor diversity-enhancing nudging design based on the accuracy motivational level (Study 2 in Chapter 4). Based on the guideline, I developed and studied three different diversity-enhancing nudging designs with social discussion platforms across two domains—politics and consumer health—and provided validation for the design guideline (Studies 2-4 in Chapters 4-6)

- **Goal 4:** *Design beyond individual with moderators* by **identifying opinion group differences in selective exposure**, in order to consider the implications for social technology designs. Beyond individuals, social science researchers have studied selective exposure at the group level—the disproportional interactions between people of the same opinion or social category, which has also been referred as *in-group bias.* In-group bias has important implications for designing diversity-enhancing social technologies, because it not only limits individuals' exposure to dissimilar views, but also impacts the social network structure, on which many design features and algorithms depend. Empirical studies, both online and offline, reported that opinion groups may differ in their selective exposure tendencies [52, 53, 54]. The potential outcome is an asymmetric social network structure that can unequally influence user behaviors, as well as the reachability, presence and impact of different viewpoints, all of which are deemed detrimental to the goal of promoting exposure to diverse views. To take the potential bias into consideration for system designs, a first step would be to identify the propensity of group difference in selective exposure, and this is where the knowledge about the moderators of selective exposure can be potentially applied. As a first exploration on the topic, I presented a case study of opinion group differences in in-group bias by analyzing large-size Twitter data, and attempted to identify association between the varied in-group bias and group attribute that is linked to moderators of selective exposure (Study 5 in Chapter 7).

Before discussing my dissertation research, in the next chapter, I will first

Table 1.1: **Overview of research goals and how they are covered in each chapter (study)**

Research goal		Chapter number				
		3	4	5	6	7
design nudges	enable negative moderators, eliminate positive ones	✓	✓	✓	✓	
design personalization	identify user groups or use contexts with varied selective exposure tendencies	✓			✓	✓
	tailor nudges for different user groups and use contexts		✓	✓	✓	
design beyond individuals	understand group differences in selective exposure and the implications for social technologies					✓

motivate my work by reviewing theoretical foundations and empirical evidence of selective exposure and its moderators, previous work on designing diversity-enhancing technologies, emerging research on studying selective exposure as a form of bias in consumer health information seeking, and selective exposure in its aggregated form—in-group bias of opinion groups on social media.

On a final note, I am by no means claiming that selective exposure is inherently wrong and that we should eradicate it under any circumstances. In fact, many would argue that, without our desire for confirmation, we would never be able to learn [55]. Selective exposure can also bring other benefits such as engagement, confidence and enthusiasm for civic life [11]. Again I emphasize the notion of a "toolbox"; as long as the needs for the tools are valid—in fact, imperative in many contexts—the tools should be developed. How to use the toolbox—when and where it should be used, how to ensure it is not abused, etc.—is an important question by its own right and should be asked by information-seekers, "choice environment" designers and policy-makers, but it is not the focus of this dissertation.

Chapter 2

RELATED WORK

In this chapter, I will review relevant work that motivated my dissertation. I will start by presenting the theoretical foundations and empirical results of research into the selective exposure phenomenon, including both social science research dated back to 1950s and recent research focusing on online selective exposure. As the main theoretical foundation of my dissertation, I will then review the psychological research on the moderators of selective exposure tendency. To motivate studying selective exposure beyond the political domain in my research, I will review the emerging research on selective exposure in consumer online health information seeking. To motivate studying selective exposure beyond the individual, in the last section, I will discuss empirical findings of in-group bias and its potential implications for diversity-enhancing technologies.

2.1 Selective Exposure to Information

Selective exposure is defined as the preferential seeking of confirmatory information, and/or avoidance or ignorance of information that challenges one's existing views. The classic view of selective exposure research is based on Festinger's cognitive dissonance theory [15], which states that individuals are motivated to preferentially seek confirmatory information to avoid cognitive dissonance—the psychological discomfort experienced from striving to maintain internal consistency when facing conflicts. Selective exposure has also been called congeniality bias [56] and a form of confirmation bias [8]; while confirmation bias also considers the preferential interpretation of information, selective exposure focuses primarily on the "exposure" stage.

While selective exposure is recognized to be a fundamental human be-

havioral tendency, it is often studied as a bias that may lead to negative outcomes. This is because selective exposure limits individuals' exposure to diverse information and impedes the many benefits that such exposure can bring. First of all, to develop accurate views and to make unbiased decisions require some degree of exposure to information that challenges one's existing beliefs. Individuals with wrong beliefs may never be able to correct them if they continue seeking confirmatory information [8]. The benefit of having diverse perspectives is also widely recognized in the literature on group decision-making [57, 58]. Second, for the individuals, exposure to diverse viewpoints is the key to preventing extremity and radicalization, and, for society, ideological polarization [13]. Third, exposure to diverse views may cultivate individuals' tolerance, empathy and acceptance of different perspectives as well the as disagreeable outcomes in the political sphere [12]. Last but not least, exposure to diverse views is the premise for "deliberative democracy," where being informed of diverse perspectives is the first and foremost step towards deliberation and democracy [9, 10, 59].

Given these benefits of exposure to diverse viewpoints for public life, perhaps not surprisingly, selective exposure has received much attention from political science and communication researchers. There has been a long-standing interest in empirically examining the extent of selective exposure as reflected in individuals' media choices. Studies consistently found that people's political preferences largely affect their choices of media sources, not only for political issues but also for "soft" subjects such as travel and entertainment [60]. For example, conservatives would consistently choose to obtain all their daily information from conservative media outlets such as Fox News. Such preferential selection happens across all media channels from radio programs to newspaper to television news [60, 61]. As a result, the choices of partisan information sources decline people's exposure to different opinions [17]. Selective exposure also happens widely in face-to-face communications. As a form of social selection [62], people preferentially choose to interact with and befriend people with whom they share similar views, even making life decisions such as where to live based on these preferences, as demonstrated in the large volume of literature on ideological segregation [63, 64].

2.2 Selective Exposure Online

In the last 15 years, a new wave of selective exposure research turned its attention to the Internet [52, 65, 66, 67, 61]. Scholars warned that there are two dangers associated with the nature of Internet media. One is the emergence of narrowcasting channels. The traditional type of broadcasting channels addresses a broad audiences, and thus has to ensure a certain level of diversity, which is further regulated by media policies [67]. In contrast, Internet users are able to choose information sources freely according to their own preferences, therefore creating self-tailored "narrowcasting" channels, as best illustrated by the hypothetical example of "Daily Me" in Negroponte's book *Being Digital* [68]. The second problem arises from the fact that popular personalization technologies often prioritize satisfying users' natural preferences, selective exposure being one of them. Examples include Google's personalized search and Facebook's personalized feed. The problem is extensively discussed in Pariser's book *The Fillter Bubble: What the Internet is Hiding from You* [19].

At a first glance, empirical evidence of online information exposure lends support to the alarming concern regarding increasing selective exposure. For example, by analyzing the URL link patterns of blogosphere, Adamic and Glance [52] found that liberals and conservatives link primarily within their ideological communities. Gilbert et al. analyzed the content of online blogs and also found that blogs tend to attract audiences that hold similar views [38]. By analyzing news consumption of users of personalized search and social media, Flaxman et al. found that they exhibit higher segregation in their news-reading behaviors than those who do not use such technologies [69].

Other studies, however, defended Internet's role in curtailing people's exposing to diverse views. The first argument is that, because of the largely decreased information access cost, although selectivity still exists, Internet is exposing people to more diverse viewpoints than traditional information sources [70, 65, 66]. There is some evidence that Internet users' information consumption may be less segregated than obtained through face-to-face communications [67] and cable news [61]. The argument is especially plausible considering the richer forms of social interactions online. For example,

Conover et al. [71] examined the Twitter network of political communications and found that, while the retweeting (i.e, sharing) network is highly partisan, the mentioning network (i.e., inter-personal interactions) is considerably heterogeneous. These observations motivated Garrett [72] to propose that individuals are driven to seek pro-attitudinal information without systematically avoiding counter-attitudinal information. Such distinction has important implications for online information exposure; the information-rich environment may provide ample opportunities for people to seek confirmatory information, and, thus, they may not necessarily need to actively filter out disagreeable information.

Another argument is that selective exposure tendency varies largely across individuals and information contexts. Therefore, the assumption that, with the diverse backgrounds and goals, Internet users will inevitably seek confirmatory information is simply problematic. Several studies reported that, when browsing online political news and discussions, some users actively seek out diverse views [37, 73], especially those with high political and topical interest [74]. Other studies reported that selective exposure tends to be weaker for certain topics [75], especially non-political ones [38]. Given that many online information sources and social media are not exclusively dedicated to a single topic, people's choices of information sources can be affected by many factors other than ideological difference alone. This motivates scholars to argue that there are many, especially non-political (e.g., hobby forums, interest-based blogs), spaces on the Internet playing an important role in people's serendipitous exposure to cross-ideological viewpoints [73, 76].

2.3 Moderators of Selective Exposure

Motivated by the empirical observation that diversity preferences differ across individuals and contexts, some scholars argue that the key to designing technologies to mitigate selective exposure is personalization [73, 77]—to leverage the various personal and contextual factors that moderate selective exposure to increase individuals' tendency to seek different opinions in specific settings. For example, Garrett and Resnick [77] proposed to 1) provide favorable challenging items according to individuals' preferences, by, e.g.,

14

tracking individual's previous activities to identify what kind of attributes makes counter-attitudinal information appealing to the particular user and 2) provide challenging items only in specific context. For example, previous research suggests that personal interest and issue importance may trump ideological difference, so we may choose to present a more diverse set of information to users for topics they are more interested in [74, 78].

While empirical studies to support this personalization approach are currently elusive, a study conducted by Munson and Resnick [73] on user interactions with a news aggregator shows that, not only are there individual differences in preference for ideologically diverse versus homogeneous news collection, but they also react differently to attitude-relevant design features. For example, those challenge-averse users, who prefer as much agreeable news as possible, were more satisfied with the system when agreeable items were highlighted. The result suggests that it may be necessary to develop different interventions for diversity-seeking and challenge-averse users, and to do so requires understanding the underlying factors that moderate individuals' selective exposure tendencies.

The contemporary discussion on the individual differences in selective exposure online is, in fact, a new wave of a longstanding inquiry. After the early critics such as Sears and Freedman [41] questioned the existence of selective exposure as an underlying psychological mechanism, selective exposure research underwent several years of decline. In 1986, Frey published a paper [79] that reviewed the accumulating mixed evidence on the existence of selective exposure, calling for research on the moderators of selective exposure to deepen our understanding on the complex mechanism for seeking and processing attitude-relevant information. The long list of moderators included: freedom of information choice [80, 81], commitment to position [82, 83], confidence [84], source competence [85], information utility [48], and decision reversibility [86]. Since then, researchers have examined many additional personal and contextual factors that make individuals more or less likely to engage in selective exposure, such as threat [49], mood [87], attitude strength [88], and framing [89, 90].

As Neuman [91] argued, the key to understanding information-seeking behavior is to understand motivations, i.e., why people seek certain types of information. By conducting a meta-analysis, Hart et al. [42] proposed the

accuracy-defense motivation framework. Embedded in theories on how different types of motivation impact the processing of attitude-relevant information [92, 93], the framework attempts to encompass the many moderators identified in the half century of selective exposure research. It provides a useful tool to interpret and predict how various individual and contextual factors moderate the selective exposure tendency.

For defense motivation, the theory resorts to the fundamental mechanism that drives selective exposure—cognitive dissonance. Therefore, factors that rise self-threat, and thus cognitive conflict, may motivate people to seek confirmatory information to defend themselves. The meta-analysis identified that selective exposure is stronger when 1) the information recipient has high commitment to the position; 2) the recipient has low confidence or perceives self-threat in an attitude, belief or behavior; 3) the recipient is close-minded; 4) the attitude is highly relevant to the personal value; and 5) no support was received compared to receiving support prior to information selection. Beyond these factors, other research consistently identified factors that give rise to self-threat, such as anxiety in information seeking [50], encountering of threatening images [94], lack of competence [95], tended to increase selective exposure tendency.

For accuracy motivation, the theory refers to motivation that promotes tendencies to seek information in an objective, open-minded fashion to foster uncovering of truth [96]. The meta-analysis mainly associated it with two factors: 1) outcome-relevant importance and 2) information utility. However, the meta-analysis also pointed out that such effects were somewhat mixed in empirical findings. The reason is that the co-existence of accuracy and defense motivation was not controlled in some empirical studies. For example, while high-quality attitude-challenging information may serve accuracy motivation, especially when the information is relevant to the current goal (e.g., expecting to write an essay addressing both sides), in some experiments it appeared to be threatening to the individuals, which also triggered defense motivation. Similarly, while outcome-relevant importance itself may reduce selective exposure tendency, in reality, it often correlates with value relevance, which could induce defense motivation when the need to defend oneself is salient (e.g., seeking information to justify one's candidate choice). The meta-analysis shows that, in some of these situations, the defense moti-

vation may even have a dominant effect.

Taking a motivational account of selective exposure allows me to make connections with the longstanding theories on the moderating role of motivation on information behaviors. One of them is the family of dual-process theories [1, 92], which predicts whether individuals' information choices are driven by content or surface features (e.g., written by a well-known author) depending on their levels of motivation to elaborately process the content. Such predictions may serve as a theoretical foundation for the provision of content or surface interface features tailored to users' motivational levels. In studies 2-4 (Chapters 4-6), I leverage this idea to develop three diversity-enhancing nudging designs. In the beginning of Chapter 4, I will discuss dual-process theories and how they may inform a theory-driven approach to personalization of diversity-enhancing technologies.

I chose to use the accuracy-defense motivation framework as the theoretical foundation for my dissertation. However, we must note the existence of other relevant theories that conceptualize the variation of selective exposure tendencies in different ways. One major thrust is to conceptualize the moderators in terms of the extent to which they regulate the emergence of cognitive dissonance (see [43]). So strategies that are able to reduce cognitive dissonance, such as providing acknowledgement of the validity of one's view, can mitigate selective exposure. It also explains past failure in observing selective exposure by insufficient generation of cognitive dissonance. To a large extent, this theme aligns with the defense-motivation process in Hart et al.'s framework.

Another line of research studies the selective exposure phenomenon through the lens of information utility (e.g., [97]). While attitude-consistent information tends to provide a higher level of gratification utility—positive mental and emotional response, attitude-consistent information would sometimes be assigned higher instrumental utility—value for attaining pragmatic goal— (e.g., when one needs to understand opponents' arguments). Across different contexts, information seekers' behaviors are governed by the single goal of maximizing information utility gain, thus exhibiting varied selective exposure tendencies depending on the utility distributions.

While this dissertation makes the choice of Hart et al.'s framework, I am

not assuming its completeness for explaining moderators of selective exposure, or its superiority over other theoretical models. The goal is to take a ground in the particular theory as an example to explore how to bridge social science theories and technology design in this area. In the last chapter, I will attempt to discuss some lessons learned on the potential limitations of the accuracy-defense-motivation framework theory.

2.4   Diversity-Enhancing Technology Design

The importance of exposure to diverse viewpoints for learning, decision-making and deliberation has engendered various technologies that take reducing selective exposure as a primary design goal. Examples of these technologies include news aggregators that aim to diversify readers' information diets, information aggregators that support decision-making or learning of controversial topics, social media for opinion exchange, and platforms for personal reflection and public deliberation.

Many of these technologies take the form of "information aggregators" with a focus on promoting awareness of diverse viewpoints. For example, NewsCube [98] is a news aggregator that classifies news for the same topic into different viewpoints and aspects and presents them in a faceted interface. Oh et al. [99] developed a blog search interface that classifies results based on their ideological leaning. Consider.It [20] is a deliberation platform where users can choose to view arguments made by people with varied positions on a controversial topic, ranging from extreme positive to extreme negative. Poli [28] is an integrated environment that aggregates comments on controversial political topics from multiple social media platforms, with the goal of enabling more effective interactions in the public sphere. Jiang et al. [100] developed a tool that aggregates diverse medical reviews from social media, and clusters them into different aspects such as effectiveness and side effects. Evaluation of these interfaces found them to be effective in facilitating learning and deliberation.

Many also explored developing visualization techniques to visually promote awareness of diverse views and to facilitate users to navigate through them [24, 25]. For example, OpinionSpace [21] is a visualization tool that projects

users' positions to points on a two-dimensional plane, where the distances between points represent their opinion differences. Reflext [101] attempts to increase users' understanding of dissimilar views by visually presenting the frequent terms appearing in comments with different stances (pro, con, neutral). DataPotrait [102] is a tool that visualizes the ideological distance and similarities on intermediary topics (e.g., hobbys) of Twitter users with the goal of encouraging exploration of diverse user profiles. Review Spoltlight [39] is a tool that supports the navigation of restaurant reviews by visually presenting adjective-noun pairs that summarize different user opinions.

As I discussed in the introduction chapter, this type of "technical HCI" research, although providing concrete design solutions, may only make limited advancement in our general knowledge about what kind of designs can effectively reduce users' selective exposure tendency. To fill this gap, we need to shift focus to identifying more general strategies that can increase users' tendency to seek attitude-challenging information. These general strategies can be considered "nudges." "Nudge" was first introduced by behavioral economists Richard Thaler and Cass Sunstein [35] as a form of libertarian paternalism—to affect behavior-change in desirable directions without interfering with people's autonomy. The core idea of nudging is the active engineering of "choice architecture," by tapping into knowledge from behavioral science on how behaviors and decisions can be influenced by the presentation of choices. Its key concept—to push individuals in the desirable directions without threatening their autonomy—is well suited to enable the success of diversity-enhancing technologies, because these technologies must be able to expose people to information they would otherwise not seek and equally importantly, to do so without significantly harming user favorability of the system and risk users abandoning the system altogether.

Based on the theory, a nudge can be an interface feature that scaffolds users' free selection among diverse information without imposing on users' choices (i.e., systems that limit information choice such as preference inconsistent recommendation [103] would not be considered a nudge, and are likely to have limitations for systems adopted by the general population). Several recent studies looked at the effectiveness of interface features as nudges for seeking diverse information. For example, Balancer [104] is a browser extension that uses an image of a person standing on a balancer to provide

feedback on the diversity (balance) of the user's news consumption, with more positive facial expression of the person corresponding to more balanced information consumption. It leverages users' normative tendency by priming diversity as social norms. Other examples include Messing and Westwood's [105] and Munson and Agapie's [106] studies, both leveraging social influences, to nudge users to attend to attitude-inconsistent information that is endorsed by other users.

To complement the technical-HCI approach of existing research in this area, my dissertation represents an effort to inform a "toolbox of nudges", from which future diversity-enhancing technologies can seek inspiration. I argue that such effort requires 1) solid evaluation that focuses on the effectiveness of reducing selective exposure per se and information-seeking outcomes and 2) seeking more generalizable conclusions, which may involve theory-based inductive and deductive reasoning by synthesizing evidence from multiple empirical studies. To this end, in Chapters 4-6, I design and study three different nudging designs with systematic evaluations of their effect on users' information selection, information reception, and learning outcomes. Taken together, these studies validate a theory-based design guideline I will discuss in the beginning of Chapter 4.

2.5   Selective Exposure in the Health Domain

While political information seeking has been the frontier for studying online selective exposure, scholars have explored the problems in other domains. The concern has mainly been in its detrimental effect on decision-making in critical contexts. Among others, an important, emerging area to study online selective exposure is consumer health information seeking.

Notably, researchers have studied selective exposure in the health domain for decades. One form of selective exposure is to seek confirmatory evidence in clinical decision-making, which has been identified to be a top cause of medical decision error [107]. For instance, when making diagnostic decisions, doctors are found to seek further evidence for presumed diagnosis but neglect or misinterpret evidence for alternative explanations. Another form of selective exposure is patients' information avoidance as a coping strategy,

especially to avoid threatening information pointing to negative outcomes [108, 109, 110]. Public health scholars have also studied selective exposure for health-risky behavior changes, as it may drive people who are inclined to engage in risky behaviors to seek evidence for justification. For example, there has been a increasing interest in studying how selective exposure mediates the effectiveness of HIV-prevention interventions [111]. Confirmation seeking from like-minded others has also been reported in recent research on the rise of online communities centered around health-risky activities, such as eating disorders and other self-harming activities [112].

A few recent studies examined how preexisting bias may affect consumers' health information seeking on the Internet [113]. Biased beliefs were found to lead to preferential seeking of confirmatory evidence. Selective exposure in consumer medical information seeking is especially dangerous considering the fact that, compared to health professionals, patients often lack necessary medical knowledge, and thus may form incorrect hypotheses to begin with, and encounter misinformation on random websites. Such biased information seeking is found to increase patients' illusory confidence in the wrong beliefs and result in incorrect or suboptimal decisions [114].

More broadly, recent research started to examine how the Internet environment may inherently shape cognitive biases in consumer health information seeking. "Cyberchodria" is a term coined to refer to consumers' escalation of concerns about common health problems after performing online information search. There are several reasons that the Internet may reinforce selective exposure towards negative information [31, 32, 115]. First of all, when it comes to health-related information, people are inclined to attend to negative information, especially when alarming terms (e.g., pain, death) or emotionally charged language (e.g., complaints in user reviews) are present. Second, popular information retrieval techniques that emphasize relevance and popularity instead of accuracy may not be suited for health-related information. For example, by ranking results based on features such as PageRank and click-through rate, the chance of an alarming result (e.g., serious diseases) showing up in the top rank may be much higher than the probability of the actual incidents [31]. While search engines are the primary tool that patients use to search for health information, they can inherently reinforce selective exposure. People may naturally use search terms that imply existing biases

(e.g., by searching "medicine side effect") and obtain information that is largely confirmatory. In fact, this problem was found to be the main cause of illusory confidence after performing online search.

These works suggest that diversity-enhancing designs should be introduced to health information systems to mitigate selective exposure. Similar to that of social and political issues, the learning process for health-related issues should ideally be accomplished in a reflective and unbiased way. Moreover, similar to political issues, health information seeking is an integrated part of many people's (e.g., those who have chronic conditions) daily information diet. If selective exposure happens on a regular basis, it could lead to profoundly biased beliefs that can be difficult to correct, which may threaten their health status.

With all the advancement in diversity-enhancing technologies in the political domain, I am inspired to seek opportunities to extend it to the development of technologies that nudge people to hear diverse opinions regarding health-related issues. In Chapter 6, I present a study where I transform the similar information aggregator interface into a tool that supports consumer medical decision and explore the extent of selective exposure tendency, and the effectiveness of nudging design, in the consumer health domain.

2.6 Group Difference in Selective Exposure (In-Group Bias)

On social media platforms, a direct outcome of selective exposure is in-group bias—individuals disproportionally connected with or interact with people who share similar views. At the network level, this leads to a segregated structure with limited cross-ideological or cross-opinion group interactions. As I have reviewed in the "selective exposure online" section, researchers have studied the problem of in-group bias (segregation) on various social media platforms (e.g., Twitter, Facebook blogs, discussion forum) for various opinion groups (e.g., affiliates of different political parties, people who hold different views on a controversial topic).

A few studies compared the group differences in segregation patterns on

the same social media platform. When analyzing the cross-ideological inter-actions on political blogs, Adamic and Glance [52] found that conservative blogs, compared to the liberal ones, had a denser pattern of segregation (more frequent interaction with other conservatives). Consistently, Conover et al. [53] found that conservatives tend to be more tightly connected and more active on Twitter than liberals. By looking at the Twitter activities for the 2012 US election, Christensen [116] found that leaders of a minority party tend to engage in more interactions, especially conversations with their supporters, compared to the major party members. For Egyptian politics, scholars looked at Secular and Islamist groups' Twitter activities and found that topical segregation and polarization emerged [117]; meanwhile, some found that the Islamist group became comparatively less active over time [118].

An interesting observation from these studies is that higher segregation and tighter in-group connection tend to correlate with higher activeness. While the causal relationships can be two ways, it is consistent with the notion that selective exposure may engender (illusory) confidence [8], topical interest and enthusiasm [11], and potentially radicalization [13].

The varied group tendency in selective exposure has several implications for biases in social technology. First, it may bias the diversity exposure of group members. This means that some users would live in a stronger echo chamber than others, and may gain more opinion or ideology reinforcement from using the platform. Second, it may "bias," i.e., create asymmetries in, the structure of social network. The better-connected groups are not only able to disseminate information more efficiently, but also, given that many algorithms use social network features, their contents are likely favored by these algorithms. One example is user recommender systems that take the number of followers into account, which may favor members of opinion groups that actively follow each other. Last but not least, it may introduce biases to social media analytics using the data. This is especially important consider-ing the growing interest in using social media data, particularly by "pooling" topic-related contents, to study and monitor social opinions in various do-mains such as political polls, consumer opinions, predicting stock market, etc. [119, 120, 121]. If an opinion group is more densely connected, and therefore having more engaging and supportive discussions, they may poten-

tially create more topical-relevant content, therefore making their opinions over-represented in the data.

All of these potential problems are counter-productive to the goal of promoting and preserving diverse viewpoints. Therefore, I argue that it is important to identify, even predict, group differences in selective exposure tendency, to develop solutions that can counterbalance the above-mentioned bias problems. Knowledge about the moderators of selective exposure may potentially inform such effort. For example, if members of an opinion group are more likely associated with certain defense or accuracy motivation-related factors, they could potentially drive the group's stronger or weaker tendency to engage in selective exposure.

Although the results in previous research into online communities are too sparse to conclude the association, supporting evidence could be found from social science research on varied in-group bias (often studied in the physical world). For example, a robust conclusion is that groups with a disadvantageous status, e.g., in a numerical minority (see review in [122]), with lower social status, or in a marginalized group[13], tend to show stronger in-group bias. Interestingly, the phenomena have been explained by the stronger sense of threat and insecurity that these group tend to face [123], consistent with the conclusion that self-threat is a direct contributor to defense motivation, thus heightening selective exposure.

If threat and other moderators of selective exposure can be used to predict group differences in in-group bias, it may open the door for more ethical designs of diversity-enhancing social technologies, which can counterbalance the potential biases of network structure. It may also inform the development of better analytical methods that take the potential group behavior differences into consideration. As the first study proposing this topic, I will explore the evidence of association between group attribute, as moderator of selective exposure, and group tendency in selective exposure, by conducting a case study of Twitter discussions on a controversial topic in Chapter 7.

2.7 Summary

While fruitful results have been produced in developing technologies that support navigation of and deliberation on diverse viewpoints, there lacks a fundamental understanding on what kind of design feature is effective in "nudging" people to seek more diverse opinions. While some propose a personalization approach to leverage the varied personal and contextual factors to make individuals more open to diverse views in specific settings, neither uniform frameworks that can guide the designs nor solid empirical implementation and evaluation are present. Meanwhile, the rich social science literature on moderators of selective exposure boasts both solid theoretical foundations and abounding empirical results that can inform the design of diversity-enhancing technologies. My dissertation aims to bridge the selective exposure theory scholarship and the HCI perspective on diversity-enhancing technologies.

After reviewing relevant literature, I choose to ground my research in the accuracy-defense-motivation framework proposed by Hart et al. [42], and explore its applications in multiple aspects of diversity-enhancing technology design, as in the four design goals I proposed in the introduction chapter. While the framework encompasses a broad range of moderators, for the scope of a dissertation, I choose to focus on a few factors that are highly relevant to interactions with information and social technologies.

For defense motivation, I choose to study a direct contributor to defense motivation—threat, as identified to increase the level of cognitive dissonance. In multiple studies, I show that design features that trigger threat can lead to more pronounced selective exposure thus should be avoided (Study 1, Chapter 3); that use contexts with higher level of threat tend to induce higher selective exposure tendency (Study 4, Chapter 6); and that user group with high self-threat attribute tends to show stronger selective exposure tendency (Study 5, Chapter 7). For accuracy motivation, I examine topical involvement as an important factor to consider for personalization designs regarding varied user groups and topical contexts, and find evidence that it leads to more balanced information seeking (Chapter 1, Goal 1). Moreover, by considering accuracy motivation as a key mediator in attitude-relevant information seeking, I propose and validate a theory-driven design guideline for

tailoring "nudges" for exposure to diverse opinions based on users' accuracy motivation level (Chapters 4-6, Goal 3).

Beyond the frequently studied political domain, I look into selective exposure tendency in consumer health information seeking (Chapter 6) and provide some insights on the similarities and differences of selective exposure between political and health opinion seeking. Beyond studying selective exposure at the individual level, I explore selective exposure as a group tendency on a social media platform (Chapter 7), contributing a first study that explores the association between a moderator of selective exposure and variance in in-group bias and its implications for designing and studying social technologies

In the remainder of the dissertation, I will present five studies. For each study, I will start by discussing the motivation, including reviewing additional literature, before presenting the study design and results. I will conclude the dissertation by summarizing what we learned from the studies; outlining future directions for the four goals of employing moderators to inform design of diversity-enhancing technologies, as stated in the introduction; and proposing advances in technology designs to serve the personal and public goal of promoting exposure to diverse opinions.

Chapter 3

THREAT AND TOPIC INVOLVEMENT ON SELECTIVE EXPOSURE

The increasing concerns on "echo chambers" and "filter bubbles" give rise to technologies with the primary goal of promoting the awareness and exchange of diverse viewpoints, for individual and collective deliberation on controversial topics [20, 26], and for decision-making in critical contexts such as healthcare [100], science [21], and finance [124]. These technologies were often evaluated by considering all users equally. However, recent research suggests that there are individual differences in the preference for seeking diverse versus agreeable information, and that the diversity-seeking and challenge-averse users tend to respond to attitude-relevant designs differently [37, 73].

While the individual difference in diversity preference has been repeatedly reported, few attempted to answer *what kind of users are more open to diverse viewpoints*. One exception is Knobloch-Westerwick and Meng' study [74]. By conducting a survey study on people's political information seeking habits, they found that cross-ideological information selection is more likely to happen among people with greater interest in politics, higher topical importance, and more accessible attitudes. We may also learn from selective exposure theories. The accuracy-defense motivation framework developed by Hart et al.[42], as I reviewed in the previous chapter, suggests that users who differ in diversity preference may vary in either their accuracy motivation or defense motivation.

In this study, I seek to validate that accuracy motivation related user attributes can mediate the diversity preference in interacting with information technologies. While several factors may contribute to accuracy motivation, I choose to look at topic involvement (interest and importance) as a general user attribute. Not only because topic involvement is frequently studied as a mediator for information behavior and attitude change [125], but also because it has important practical implications for customization and person-

27

alization of technologies, as many differences in user groups or use contexts can be attributed to differences in the involvement level.

Meanwhile, the empirical evidence on how topical involvement actually mediates selective exposure is somewhat mixed. To reconcile it, Hart et al. [42] suggest that, while outcome relevant involvement is supposed to contribute to accuracy motivation, in reality, it often correlates with value relevant involvement, which may conversely increase selective exposure, especially when the needs to defend one's value becomes salient. However, which type of involvement would have a dominant effect may depend on the context. Hence, empirical examination on the effect of topic involvement in mediating selective exposure in the context of interacting with information aggregators is merited.

It is worth noting that the regular use of this type of tools may not be as strongly goal-oriented as the "post-decision information gathering" paradigm used in most lab studies of selective exposure research. Thus the needs to defend value-relevant position could be relatively lower. Another argument for high topic involvement leading to more diverse exposure is that, for familiar topics, topic involvement often naturally correlates with topical knowledge and confidence, which were shown to alleviate the inclination for defense [79, 44]. It is worth pointing out, however, that increasing exposure to attitude-challenging information does not necessarily indicate compliance with the competing views. To the contrary, with topics one cares deeply about, one may scrutinize the information more elaborately, and thus may not be persuaded in the absence of strong, persuasive arguments [1, 125].

For defense motivation, which increases selective exposure, a question directly relevant to technology design is, *what factor may trigger users to be defensive and thus should be avoided?* Fundamentally, defense motivation is associated with self-threat, which increases the level of cognitive discomfort, leading to heightened motivation to seek confirmatory information to cope with it. A substantial amount of studies demonstrated that threat can exacerbate selective exposure tendency [51, 49]. Such threat can take various formats, ranging from threatening texts [43, 87], threatening images [94], task anxiety [51], challenging or negative feedback [126], information scarcity [49], etc.

Meanwhile, I note that many design features can be associated with these formats of threat, and thus potentially heighten the perceived threat in users' information-seeking context. Examples include features (e.g., images, alarming texts, etc.) that prime users to consider the negative outcomes of the topic, or information presentations that create the impression of information scarcity or lack of support. These features are pervasive. Some may be intentionally created with the expectation to make users more vigilant in their information seeking process. Others are unintentionally produced as decorative features. In this study, I examine how design features that prime threat impact users' selective exposure, and if so, to present a warning to avoid features underlining factors that may increase selective exposure.

3.1   Overview of the Study

In this work, I studied users' selective exposure using an information aggregator that presented diverse viewpoints for controversial social-political topics. I examined how an individual factor—topical involvement, and a design feature—images that primed threat, impacted users' selective exposure. I attempted to bridge the theoretical framework on moderators of selective exposure and diversity-enhancing technology design, by leveraging individual and contextual factors that could change an individual's selective exposure tendency.

Beyond information selection, I also examined how the two factors—primed threat and topic involvement—impacted information reception and information seeking outcomes. Specifically, how they influenced users' agreement with different opinions, and their attitude after exposure to diverse viewpoints. I list the research questions and main findings from the study in Table 3.1.

3.2   Methodology

In this section, I will first introduce the prototype platform a research assistant and I developed and the content material for the experiment. I will

Table 3.1: **Research questions and findings in Study 1 (Chapter 3)**

	Research question	**Findings**
RQ1	How does design feature priming threat influence users' selective exposure?	Design feature that primes threat increases users' tendency to engage in selective exposure.
RQ2	How does topic involvement impact users' selective exposure?	Topic involvement moderates the impact of threat, leading to more balanced information seeking regardless of the contextual threat.
RQ3	How does primed threat and/or topic involvement impact users' agreement with information of varied positions?	Topic involvement leads to lower agreeableness with information supporting the opposite position.
RQ4	How does primed threat and/or topic involvement impact users' attitude change after exposure to diverse viewpoints?	Both topic involvement and primed threat decrease attitude change, showing that attitude change is impacted, but not only by information exposure.

then discuss the experimental design, measurements used, participants, and procedure.

3.2.1 Platform

We built an information aggregator prototype that presents diverse viewpoints for controversial topics. For each topic, the system presented both factual arguments and social opinions (i.e. peer user comments) from both sides (pro v.s. con). The interface is shown in Figure 3.1. The web page presented factual arguments on the top half and user opinions at the bottom half. Both were presented in a two-column table, with all supportive messages on one side and opposing ones on the other. This two-column format is widely used by diversity-enhancing technologies (e.g., [26, 20]), and was found to encourage more balanced information seeking than interfaces that mix all messages in one column [26].

Each participant was asked to complete 8 tasks, each for a controversial topic. For each topic, 8 pros and cons arguments, and 5 pros and cons user comments were shown. The order of them was randomized. For each
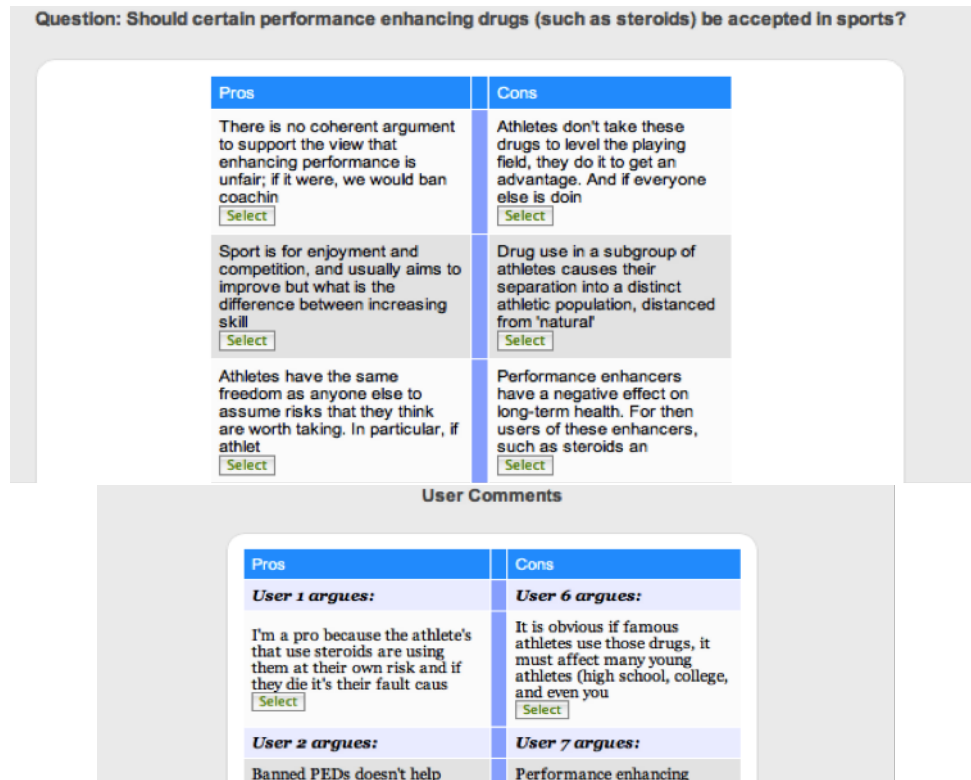
Figure 3.1: **Experiment interface for Study 1 (Chapter 3): sections presenting factual arguments (top) and user opinions (bottom)**

argument or comment, a snippet was shown to give participants a summary of the message. If the participant chose to click on "read more" for a snippet, then a pop-up window would show up with the complete message, where participants would be asked to rate their agreement with the message based on a 5-point Likert scale.

3.2.2 Experiment Material

I started by selecting 13 candidate topics that are commonly deemed as controversial (see Appendix A). To ensure a reasonably balanced distribution of topic involvement among the participants, I selected topics in various domains (e.g., ethics, healthcare, crime, sports), ranging from common focus of public debate (e.g., death penalty) to less discussed topics (e.g., using steroids for sports). A research assistant and I manually selected items of arguments and user comments on each topic from the website `www.procon.org`.

31

The website is maintained by a non-profit organization with the goal of providing resources for critical thinking for various controversial topics. For each controversial topic, the website provides arguments on both sides by summarizing information from multiple formal sources, including academic publications, newspaper, government documents, etc. For example, a supportive argument for the question "*does violent video games contribute to increases of youth violence?*" is:

> "*Violent video games desensitize players to real-life violence. It is common for victims in video games to disappear off screen when they are killed or for players to have multiple lives. In a 2005 study, violent video game exposure has been linked to reduced P300 amplitudes in the brain, which is associated with desensitization to violence and increases in aggressive behavior.*"

The website also has a "user comment" section for users to post their opinions. An example of a supportive user comment for the same topic is:

> "*Violence influences the mind, brain, and the way we act on what we would've just seen. Those thoughts would still be in our mind even after an hour or so because our mind is still re-playing what we saw on screen. This would then reflect on our actions and how we think for 30-45 minutes. Even I have experienced this.*"

We randomly selected factual arguments and user comments from the website, and made minor revision to the material to ensure that there was no significant difference in length (about 60-100 words) or rigor of arguments between items on each side. We also rewrote the first sentence, where it was necessary, to provide a summary of the message, which was shown on the front page interface as the message snippet.

3.2.3   Experimental Design

To study the effect of design feature that primed perceived threat, a repeated-measure between-subject experimental design was used. That is, half of the participants were assigned to the condition with an image feature that depicts threatening outcomes relevant to the topic. Each participant completed six tasks. A similar approach was used by Fisher et al. [49] in their serial

studies looking at the effect of threat on selective exposure. To avoid having pictures serving as biased arguments, we chose ones that highlighted negative outcomes related to the topic, but remained neutral in terms of the discussion itself. For example, for the question *"do violent video games contribute to increases of youth violence"*, we used the image shown in Figure 3.2, depicting a badly injured teenager. The image would highlight the threatening aspect of youth violence without suggesting whether violent video game was the cause.

Figure 3.2: **Image used to prime threat for the controversial topic "do violent video games contribute to increases of youth violence?"**

3.2.4 Measurements

Attitude Index

For each of the initial 13 topics, I measured the participants' prior attitudes on the topic using a 5-item semantic differential scale, which is often used to derive attitude towards a given concept by measuring its connotative meanings [127]. For example, when measuring the participants' attitude on the issue of vegetarianism, instead of directly asking whether they held a positive or negative attitude, I asked them to choose their position on a 7-point Likert scale for five pairs of bipolar adjectives: unfavorable-favorable, bad-good, unnecessary-necessary, harmful-beneficial, unhealthy-healthy. I calculated

the mean rating of the five items as the subject's prior *attitude index* on the topic. An example of the attitude measurement questionnaire used in the experiment can be found in Appendix B. The Cronbach's alpha for of the 5-item scale for all topics reached 0.87, which was close to excellent internal consistency [128]. After each experiment task, the same set of questions was asked again to measure the participant's attitude after using the tool.

Topic Involvement

Given that previous research suggested differentiating between value-relevant involvement and outcome-relevant involvement [93, 42], I started by measuring both by following Johnson and Eagly's definitions [93] (see Appendix B):

1) Value-relevant involvement refers to the extent to which the topic is linked to important personal value. I measured value-relevant involvement by asking participants (all based on a 7-point Likert scale):

    a) *How much is this topic related to your core value?*

    b) *How important is your point of view on this issue to you?*

2) Outcome-relevant involvement refers to the extent to which one is motivated to learn about the topic to ensure correct understanding of the topic. I measured outcome-relevant involvement by asking:

    a) *How interested are you in learning more about the topic?*

    b) *How much do you desire to know the truth about the topic, regardless of your own point of view?*

It turned out the results of the two types of involvement measures were highly correlated ($r = 0.84$), which echoed the conclusion from previous studies on topic involvement [42, 125]. I therefore chose to combine the two measures to create a single index for topical involvement by averaging the ratings of the four questions above. The Cronbach's alpha of the four items reached 0.93, which is considered excellent internal consistency in measuring the same latent variable of topic involvement.

3.2.5 Participants

28 participants were recruited from the Champaign-Urbana area in Illinois, USA. They mainly consisted of a mix of college students, university staff, and employees of other industries. According to Dutta-Bergman [129], younger people with higher education are most likely to seek political news and participate in political activities online. Although it is possible that people with different social, cultural, and educational backgrounds may exhibit different information seeking behaviors, the sample was representative of the most active group who seek information about social-political issues online.

Participants were randomly assigned to the condition with (Group 2) and without (Group 1) images that primed the perceived threat. The demographic questionnaire showed that there was no significant difference ($p > 0.10$) in age ($M_1 = 26.30, M_2 = 28.78$), gender (64.3% in group 1 and 57.1% in group 2 are female) education level (35.7% in group 1 and 42.9% in group 2 are graduate students or have graduate degree), Internet use frequency ($M_1 = 1.72, M_2 = 1.53$, for scale from 1-less than an hour per day to 5-more than 8 hours per day), political leaning ($M_1 = 3.4, M_2 = 3.1$, scale from 1-conservative to 5-liberal), and self-reported knowledge about the topics between the two groups.

3.2.6 Procedure

One week before the experiment, participants were contacted by email to complete an online survey to measure their prior attitudes and topic involvement for all the 13 candidate topics, as well as demographic information. After that, I excluded topics that were highly imbalanced in the pre-existing attitudes and topic involvement scales. Specifically, I excluded topics in which the number of people on one side of the attitude or topic involvement scale were more than two times the number of people on the other side. This was done to ensure that for each topic there was a balanced number of participants having high or low pre-existing attitudes, and high or low involvement. This left us with 8 topics used in the experiment. Examples included "should euthanasia be legal?" and "should people become vegetarian?" The complete topic list can be found in the Appendix A.

Participants were invited to the lab to complete the experiment. For each of the 8 tasks, participants assigned to the experimental condition were firstly presented with the threat-priming image, as described in the previous section, while participants in the control condition skipped this step. Then they proceeded to use the information aggregator. They were asked to freely browse the information, and afterwards to write a short message about their own view on the topic. Participants were told to read any number of items in any order, and there was no time limit for browsing. The system automatically recorded their interactions for analysis. At the end of each task, participants were asked to finish a questionnaire to measure their attitude on the topic again.

3.3   Results

To differentiate between topics that participants had low or high involvement in, I performed median splits on the topic involvement index of the eight topics for each participant. That is, each participant had 4 high-involvement and 4 low-involvement topics. The mean value of the medians for all the participants was 3.75 ($SD = 0.60$), and there was no outlier (more than +/- 3 SD of the mean). It means that the levels of topic involvement were well distributed across the participants.

I recoded the participants' prior attitude as positive or negative according to whether the prior attitude index was more or less than the median (rating=4), which was also the median value of attitude ratings from all the participants. For 6 out of the 224 cases where the participants gave the neutral rating 4, I randomly assigned them to either side. I examined the proportion of tasks with positive attitudes for the two experiment groups (threat and no-threat), and found they did not significantly differ in their prior attitudes ($\chi^2(1, 224) = 0.18, p = 0.89$).

In this study, I did not control for the magnitude of the prior attitude bias during sampling. To exclude the potential influence of attitude extremity on interpreting the potential effect, I examined whether there was correlation between topic involvement index and attitude extremity, defined as the absolute deviation of prior attitude rating from neutral (rating=4), and found

no significant correlation($r = -0.06$). I also examined the effect of attitude extremity on selective exposure but found no significant effect.

On average, participants selected (i.e., clicked on) 8.28 items (SD=4.93), including 6.12 (SD=3.64) arguments and 2.16 (SD=2.16) comments, and they spent M=33.98 (SD=30.39) seconds reading each item. There was no significant difference between the two experiment groups in all these measures described above, suggesting that the general participation level was about the same between the two groups.

In the rest of this section, I will first discuss how primed threat and topic involvement mediated participants' selective exposure tendency to answer RQ1 and RQ2. I will then examine how the two factors impacted participants' agreement with information regarding the two competing sides (RQ3), and their attitude changes (RQ4), by comparing the attitude index measured in the post-study and prior-study questionnaires.

3.3.1   Selective Exposure (RQ1 and RQ2)

To examine participants' exposure to attitude consistent and inconsistent information, I created a *Selective Exposure Index (SEI)*, by calculating the difference between the number of attitude-consistent and attitude-inconsistent messages clicked for each topic. Here whether a message was attitude consistent or inconsistent was decided by whether the message position was consistent with the coded prior attitude index (positive/negative) of the participant, as described above. That is, if a participant held a positive prior attitude on the topic, then a pro argument or comment was considered attitude-consistent while a con argument or comment was considered attitude-inconsistent. The SEI was calculated to reflect the extent of selective exposure, such that if it was a positive value, it meant selective exposure presented, and a higher magnitude indicated a stronger selective exposure bias.

I performed a two-way repeated measure ANOVA on SEI with primed threat (without/with threatening image) as between-subjects variable, and topic involvement (low/high) as within-subjects variable. The result showed that the main effect of threat was significant ($F(1, 26) = 9.61, p < 0.01$),

37

and there was a significant two-way interaction between primed threat and topic involvement $(F(1,26) = 4.91, p = 0.04)$. Figure 3.3 plots the mean value of SEI for different conditions. The figure shows that the image feature priming topical threat provoked stronger selective exposure, but only for low-involvement topics. It is further confirmed by the significant main effect of the presence of threat-priming feature on SEI for topics one had low involvement $(F(1,26) = 10.46, p < 0.01)$. In contrast, the feature had little effect on topics participants had high involvement, as they behaved consistently by seeking balanced information regardless of the contextual threat $(F(1,26) = 0.02, p = 0.90)$.
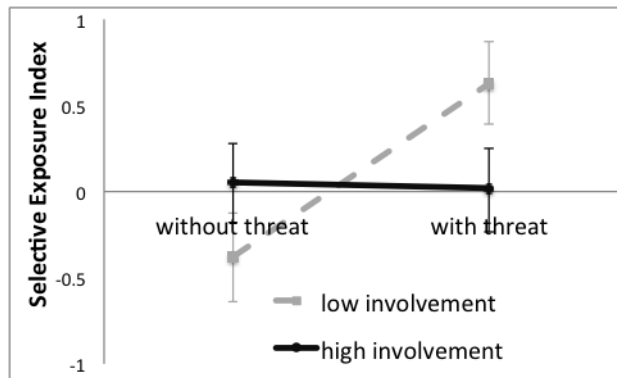
Figure 3.3: **Average SEI for topics one had low/high involvement, with/without threat-priming feature**

In addition, I performed one sample t-tests to compare each participant's average SEI to zero (balanced in selecting both sides) with each of the four combinations of with/without threat and high/low topic involvement. The result showed that none were significantly different from zero, except when participants were presented with threat-priming images for topics they had low involvement,$(t(13) = 2.57, p = 0.02)$. The above results showed that *when participants were presented with design features that prime contextually relevant threat, they exhibited pronounced selective exposure for topics that they had low involvement, but not for topics they had high involvement.*

First of all, the results provided validation that threat as a moderator contributing to defense motivation increases selective exposure tendency. The result was consistent with conclusions from multiple previous studies demonstrating the positive moderating effect of threat on selective exposure [49].

I also found that topic involvement, which is considered a main contributor of accuracy motivation, could moderate such effect and lead to balanced information-seeking regardless of the contextual threat. While I did not observe its moderating effect on selective exposure in the control condition, it might be attributed to the fact that the baseline selective exposure was already close to none in the use of the two-column interface [99]. Still, the results could be considered evidence that supports the theoretical prediction that defense motivation is a positive moderator, and accuracy motivation is a negative moderator, of selective exposure. The results also further suggested that there could be interaction between the two types of motivation. Specifically, individuals' high topic involvement could make them less likely to be impacted by factors that triggered defense motivation.

3.3.2   Information Judgment (RQ2)

Users' attitude change after being exposed to an information-diverse environment is likely not only influenced by the information exposure, but also by their reception of the information, such as agreement with the arguments. To analyze the participants' agreement with attitude consistent and attitude inconsistent information, I created another variable—*selective rating*, by calculating the difference between the average rating given to all attitude consistent items and that of attitude inconsistent items for each topic. A positive selective rating would indicate that attitude consistent information was evaluated more favorably, and a higher magnitude meant this preference was stronger.

I performed a two-way repeated measure ANOVA on selective rating with perceived threat (with/without images priming threat) as between-subjects variable and topic involvement (high/low) as within-subjects variable. I found that the main effect of topic involvement was significant ($F(1, 28) = 4.24 p = 0.05$). No effect of threat-priming features was observed. As illustrated in Figure 3.4, *participants showed stronger disagreement with attitude-inconsistent information for topics they had high involvement.* The result echoed conclusions from previous studies, suggesting that high topic involvement could promote the tendency to critically scrutinize the information

content [1, 125]
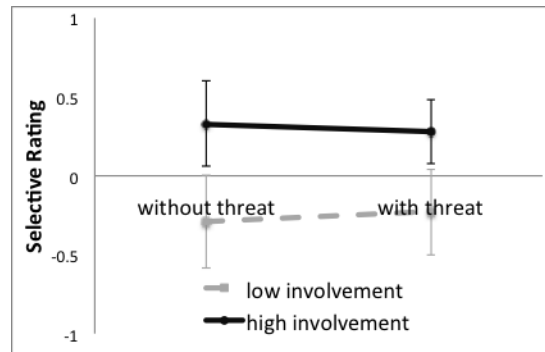
Figure 3.4: **Average Selective Ratings for topics one had low/high involvement, with/without threat-priming features**

3.3.3 Attitude Change (RQ3)

An important goal of exposing people to diverse viewpoints is to prevent attitude polarization. Therefore, when examining attitude change, I conceptually distinguished betweeb situations where attitude was moderated, i.e., moved to the opposite direction of one's prior attitude, and situations where attitude became more extreme, i.e., moved further away in the same direction of the prior attitude. Given that the prior and post attitude measurement was based on a 1 (negative) to 7 (positive) scale, I created an *attitude moderation index* by: 1) If the participant held a positive prior attitude, attitude moderation was calculated by the difference of prior attitude index and post attitude index; 2) If the participant held a negative prior attitude, attitude moderation was calculated by the difference of post attitude index and prior attitude index. Hence a positive attitude moderation value would indicate that the participant's attitude was moderated, while a negative value would indicate that it became more extreme, and the magnitude of the index indicated the extent of attitude change in either direction.

In the experiment, both for topics with high $(t(29) = 7.15., p < 0.01)$ and low involvement $(t(29) = 7.89, p < 0.01)$, the participants' average attitude moderation was significantly higher than zero (no change), suggesting that for both types of topics individual's attitudes were moderated after having

used the system presenting diverse information.

To examine how topic involvement and primed threat impacted attitude moderation, I first performed a two-way repeated measure ANOVA on attitude moderation index with primed threat (with/without threat) as between-subjects variable and topic involvement as within-subjects variable. I found that the main effect of topic involvement was significant ($F(1, 26) = 12.82, p < 0.01$), and the main effect of the presence of threat was marginally significant ($F(1, 26) = 3.09, p = 0.09$). It suggested that, as shown in Figure 3.5, the participants' attitudes were less subject to moderation for topics they had high involvement, and the presence of threat also led to slightly less moderation of attitude.
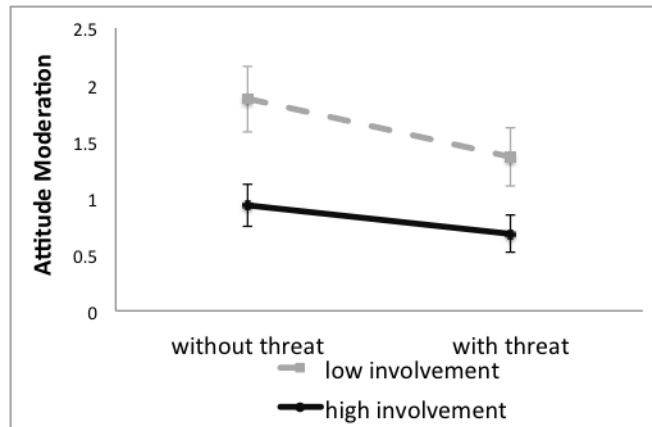
Figure 3.5: **Average Attitude Moderation Index for topics one had low/high involvement, with/without threat-priming feature**

Considering the results in the information selection, I argue that attitude moderation is not only impacted by information exposure, but also other factors such as scrutiny of information content, and attitude strength. On the one hand, by increasing selective exposure, primed contextual threat lessened exposure to information that challenges one's existing attitude, and thus the opportunities for attitude moderation. On the other hand, although people who had high topic involvement tended to engage in a more balanced information-seeking pattern, they were less likely to be persuaded to change their attitudes. One reason, as shown in the information judgment section, could be that participants were more critical of attitude-inconsistent information, hence more resistant to attitude moderation in the absence of strong,

41

meaningful arguments. Moreover, topic involvement is often found to be correlated with topical knowledge, attitude strength and attitude confidence, and thus naturally it would be more difficult to change their attitudes, even in the presence of arguments to correct inaccurate beliefs.

3.4 Discussion

First of all, this study presents a warning for technologies that host diverse viewpoints to avoid factors, including design features, that may provoke defense motivation and cognitive discomfort in users. This may include avoiding threatening texts and images, anxiety-inducing factors such as time pressure and information scarcity, and features that accentuate challenges or threaten self-confidence, etc. This is, in fact, the core notion underlying many design solutions proposed to reduce selective exposure, such as by making supportive information easily accessible [77], and by partially acknowledging users' position before introducing "de-biasing" information in decision-support systems [130].

By identifying threat as a positive moderator of selective exposure in using information technologies, the results may also inform tailoring designs for different user groups and use contexts. It is possible that user groups facing a higher level of threat, such as those who are frequently challenged, may have a stronger selective exposure tendency. Also, a person who is open-minded in regular social-opinion seeking activities may potentially become more challenge-averse when it is for acute information needs (e.g., learning about ongoing events, making time sensitive decisions), or when it is about a more threatening topic (e.g., finance, health). In Chapter 6 and Chapter 7 (study 4 and study 5), I will further explore whether contextual threat leads to more pronounced selective exposure in certain domains (e.g., health threats), and for certain user groups (e.g., minority opinion group).

Second, the study suggests that topic involvement is an explanatory factor for the user difference in diversity-seeking versus challenge-averse tendencies observed in previous studies [73], and echoed the observation in Knobloch-Westerwick and Meng' survey study [74]. From a personalization perspective, the study suggests that for topics that users are highly interested in or con-

sider to be important, they may be more open to, or even favor, information collections that provide a diverse set of viewpoints. However, for topics that users are less interested in or consider unimportant, they may be more prone to selective exposure. In such situations, one should carefully tailor the information retrieval and presentation to ensure exposure to diverse views without significantly harming the user opinion of the system.

The association between topic involvement and diversity-seeking could be bad news. It implies that people who are less involved in a topic, therefore likely less knowledgeable and less frequently exposed to the topic, are more likely to engage in selective exposure, which could potentially deepen their "uninformed bias". The correlation between topic involvement and topical knowledge was confirmed in our data with $r = 0.77$. A possibly more alarming observation is that there is a low correlation between attitude extremity (measured by absolute difference between prior attitude rating and neutral rating 4) and topic involvement ($r = -0.06$). In other words, some people may hold a relatively extreme attitude even though they have neither adequate knowledge nor motivation to learn about the topic. Exposing these people to diverse viewpoints is particularly necessary, but challenging given that they are more inclined to engage in selective exposure, even in our experiment with an already optimized interface design that presents competing views side by side. In Chapter 5, I will explore how to design to nudge people with low accuracy motivation to seek high-quality diverse views.

The study paints an interesting picture for users who have high topic involvement—those who consider a topic to be important and are eager to learn, meanwhile also tend to be relatively knowledgeable. On the one hand, they are inclined to seek information in a more balanced fashion regardless of the contexts. On the other hand, they also appear to be more critical in judging the content, and are less likely to be persuaded in the absence of meaningful, persuasive arguments. Hence, just because they are less likely to engage in selective exposure, it does not mean they do not need to be nudged towards diverse opinions. On the one hand, in the presence of biased views or incorrect beliefs, they may be the group in need of more persistent de-biasing effort. On the other hand, designing diversity-enhancing technologies for this group of users should emphasize the access to diverse perspectives as an inherent user needs. This calls for a more fine-grained

view on categorizing attitude-challenging information, by identifying what kind of attitude-challenging information would be favored for facilitating the learning process. In Chapter 4 and Chapter 6, I will explore some answers.

Chapter 4

STUDYING POSITION INDICATORS AS A CENTRAL NUDGE

In the first study, I found evidence that accuracy motivation related factors can mediate how people seek attitude relevant information. Now I ask, how does accuracy motivation mediate the effectiveness of designs that nudge users to seek diverse views? That is, how should we tailor nudging designs for users with high or low accuracy motivation?

In fact, accuracy motivation is considered a major type of motivational factor mediating information seeking and processing [131], and has been studied in the large volume of literature on dual-process theories. Dual process theories have their roots in the assumption that cognitive processes can be divided into two general classes: those operating automatically, and those operating in a controlled fashion [132, 133]. The automatic system is often elicited unintentionally and uses little cognitive resource to process information. The controlled system operates with considerable amount of cognitive resources and is often initiated intentionally.

Dual process theories gained prominence in studying persuasion, as in explaining how persuasive messages are processed by individuals to change their attitudes and behaviors. Two of the most influential dual-process persuasion theories are Elaboration Likelihood Model (ELM; [1]) and Heuristic Systematic Model (HSM; [134]). The central notion of ELM is that individuals' extent of elaboration to process persuasive messages (i.e., engaging in controlled instead of automatic processing) is determined by two factors—motivation and ability (as illustrated in Figure 4.1). In other words, when the motivation and ability is high, individuals engage in careful scrutiny of, and deliberation on, the information content, in order to obtain a carefully considered evaluation. This is often called the central route of information processing. Conversely, when either the motivation or ability is low, people engage in the peripheral route of information processing with low level of
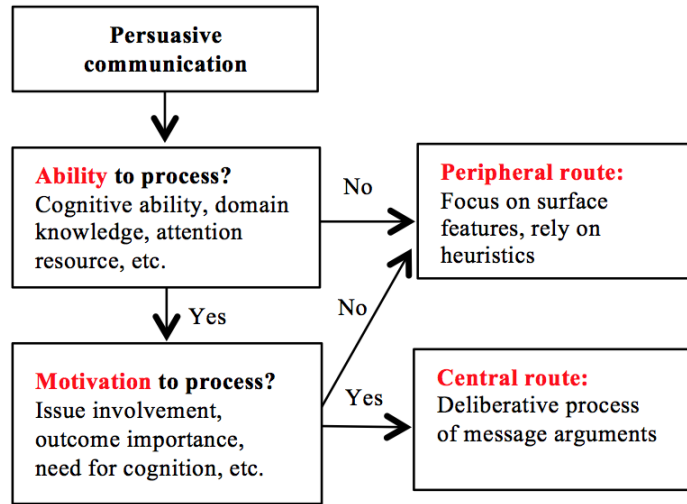
Figure 4.1: **Elaboration Likelihood Model [1]**

elaboration. Through the peripheral route, attitude-change can be affected by the use of heuristics (e.g., experts say it right, or I agree with what many others agree), or other types of information processing shortcuts. Similarly, HSM describes the two types of persuasion processes as systematic processing and heuristic processing. A difference between HSM and ELM is that HSM deems the two processes as separate, but can happen simultaneously, while ELM considers elaboration as a continuum and study the extent of elaboration (i.e. extent of central-route processing). Dual process models are often used to tailor persuasion strategies according to the motivational and capability factors of the audience and the information-delivery context. For example, for TV and billboard advertisement, since people are unlikely to have high motivation or time (capability) to process the information content, peripheral-route is often targeted by, for instance, intriguing pictures or attractive endorsers.

The concept of the dual cognitive processes is also a core notion in the "nudge" literature. Sunstein [135] differentiates between two classes of nudges: nudges that target the central processing system by promoting conscious deliberation, and nudges that target the peripheral processing system by enlisting automatic behavioral tendencies. For example, to nudge people into healthier eating habits, Sunstein considers nutrition labels as an example in

46

the first category, as in nudging people to make conscious healthy choices. In contrast, the strategy used by some cafeteria to put healthy food in the front row is considered an example in the second category, as in taking advantage of people's automatic behavioral tendencies that can be out of their conscious awareness.

The review of the dual-processing persuasion models suggest that motivational and capability factors may mediate the effectiveness of these nudges. For example, nutrition labels may enhance the likelihood of buying healthier food for those who are motivated to live healthy and are able to understand the labels, but may not work for those who do not care about the consequences or simply cannot interpret the nutrition information. Therefore, it may be necessary to tailor nudging design based on users' motivational and capability levels.

Such a view resonates with recent advocates in the field of persuasive technology [136] for personalization designs. Some argue for differentiating between, and applying different persuasion strategies to individuals who differ in their susceptibility to the targeted behavior changes [137]. Some suggest differentiating between users' motivational types to apply different rewarding mechanisms for behavior changes [138]. Others attempt to classify persuasive designs into different categories and map them to different user groups or contexts. In particular, to design "nudges" that can improve users' privacy decisions, which is one of the most researched topics in nudging design for technologies, several studies [139, 140, 141, 142] have applied the ELM theory to reconcile people's varying tendencies to engage in rational privacy-related behaviors (i.e., "privacy calculus"; [143]) and to rely on heuristic judgments (e.g., social proof, affect). To tailor nudging designs, these studies recommend providing rich information to support rational decision-making for those who have high privacy concerns (i.e., motivation), those who are knowledgeable about privacy decisions, and those that are more capable of coping with the decisions (i.e., capability). In contrast, to nudge people who do not have these attributes, these studies suggest providing "heuristic cues" such as third-party endorsement and reputation scores.

4.1 Designing Personalized Diversity-Enhancing Nudges: An ELM-Based Approach

In a series of studies in the following chapters (Study 2-4, Chapter 4-6), I present an ELM based design guideline that tailors "nudging cues" for diversity-enhancing information systems based on individuals' accuracy motivation level. It means that I address three goals. First of all, I study what kind of information cues may increase users' tendency to seek attitude-challenging information. By "information cues", I adopt a definition similar to "information scent" in information foraging theory (IFT) [144]—*proximal cues that bring about (imperfect) perception of the value, cost, or access path of information.* The scope of information cues can be broad, ranging from common interface cues such as snippets, links, icons, etc., to more sophisticated design features such as tag clouds, interface structure, visualization [145], as long as they intend to reveal information characteristics and can guide the navigation behavior.

In this work, I study the proximal cues of information as "nudges" to push users towards more desirable information-seeking path—exposure to more diverse viewpoints. I chose to study icon features that have become commonplace for contemporary information and social technologies. Examples include cues indicating the topical focuses (e.g., tags), social endorsement (e.g., likes, thumbs-up), and source related features (e.g., citation source, "expert badges" for authors). All of these features can influence users' choices of information and navigation paths. However, by studying a basic format of these cues, my interest is beyond a specific design but on a category of cues that reflect the particular characteristic being studied. For example, while I study a simple tag that reflects on which topical aspect the information focuses on, I would expect that the conclusion can inform a broad range of designs that intend to show the topical focuses of the information, such as faceted interface.

Second, based on ELM, I differentiate between two types of nudges: 1) Central nudging cues that reflect the characteristics of the *information content*, and facilitate *conscious choices* of information for elaborate processing. For example, the frequently studied link text and snippet, based on which users can choose to further explore the information in the linked page, would

fall into this category. In my work, I explore what kind of central nudges, i.e., what content related features of attitude-inconsistent information would make it more appealing for information seekers. 2) Peripheral nudging cues that reflect the *"peripheral features"* (i.e., not integral part of the content) of the information, and induce *heuristics* that can affect users' relatively automatic choices of information. For example, source related features often invoke source related heuristics such as "experts say it right". The frequently encountered social endorsement features, such as likes, sharing, etc., are often associated with social proof related heuristics. In my work, I look beyond whether peripheral nudges can influence users' information selection, which has already been repeatedly manifested in previous work [146, 147], but explore what features may interact with information position, e.g., nudging *more* for seeking attitude-inconsistent than consistent information, and thus reducing selective exposure to (at least certain types of) information (Study 3, Chapter 5).

Lastly, I study how accuracy motivation mediates the effectiveness of central nudges and peripheral nudges. Based on ELM, I hypothesize that central nudges would be more effective in reducing selective exposure for those with high accuracy motivation than those with low accuracy motivation. The difference of effectiveness between the two groups may be lower, or even reverse, for peripheral nudges. I argue that, to design technologies for behavior changes, it may be necessary to differentiate between those who need to be *supported*, and those who need to be *transformed*. This is relevant to the notion of differentiating between "means oriented nudge" and "ends oriented nudge". While the former targets providing support for people to attain their own goals, the latter works by adjusting people's behavioral goals. In this sense, understanding what nudges are effective in facilitating those who have high accuracy motivation to better explore diverse viewpoints is not only to serve the diversity-enhancing goal, but also to satisfy the inherent user needs.

4.2   Overview of the Chapter

As the first of the three serial studies, I studied the effect of a central nudge—position indicators, which showed the fine-grained position reflected in the

message content, ranging from extreme to moderate. It is a central nudge because it reflects the characteristics of the information content and is expected to facilitate the careful exploration of arguments made by people with varied positions. In Chapter 5 and Chapter 6, I study a peripheral nudge (expertise indicator), and another central nudge (aspect indicator).

From a cognitive perspective, selective exposure can be attributed to people's tendency to preserve cognitive resources, as processing and counter-arguing about disagreeable information is more cognitively demanding [49, 148]. Providing explicit labels of stance may reduce the cognitive demand for processes to interpret and evaluate the message position, and thus is expected to reduce the extent of selective exposure. However, previous studies presented mixed results regarding the effectiveness of labeling information stance on reducing selective exposure [73, 99].

In this study, I further improved the way to indicate the stance of information. Most systems presenting diverse views use a dichotomous model of positions, by either a side-to-side presentation [20, 100] or labeling information as pro v.s. against, or conservative v.s. liberal [73, 99, 149]. However, I argue that using more fine-grained categories has several benefits for supporting exploration and deliberation. First, it is more ecologically valid, as people's positions naturally align on a continuum instead of dichotomy. Second, people who take more moderate positions may offer fundamentally different arguments from those with extreme positions. Understanding the diverse reasons for people to hold different views, instead of seeing them as a unitary group, is a key aspect of public deliberation. Lastly, and perhaps most importantly, distinguishing positions that are only moderately different may reduce people's resistance to attitude-challenging information. There is substantial evidence that lower attitude discrepancy between the message's and the recipient's position may reduce cognitive dissonance [150]. A few studies incorporated similar ideas in system design. For example, Opinion-Space developed by Faridani et al. [21] uses spatial distance to represent the magnitude of opinion differences. However, the study focused on examining the overall user experience without studying how features about position magnitude impacted users' information seeking process.

Since position indicators is considered a type of central nudge by reflecting the content characteristic, I hypothesized that it would be more effective

in increasing exposure to diverse views for people with high level of accuracy motivation. For these people, providing explicit labels of stances could scaffold the information seeking process in the direction they desire for, i.e., developing more accurate views. However, for people with low accuracy motivation thus less likely to make deliberative choices of information, their information behaviors would be less affected by content related features.

Beyond information selection, I also expected position indicators to impact users' information reception and learning outcomes. For information reception, I examined how the indicators induced changes in users' perception of the positions of, and their agreement with the information. For learning outcomes, I was interested in whether the position indicators could facilitate users to acquire new knowledge about the topic, especially knowledge regarding the competing view. Incorrect beliefs and biased decisions are often caused by false perception of consensus [151] and ignorance of facts that may support the alternative views [152]. Therefore, the measurements of selection, position judgment, agreement, and knowledge gain would be directly relevant for the goal of designing technological interventions that support meaningful deliberation and potentially correct biased views. In summary, I asked the following research questions listed in Table 4.1.

4.3   Methodology

In this section, I will first introduce the platform and the content material for conducting the experiments in this and the next chapter. I will then discuss the experimental design, procedure, participants and measurements used in the experiment.

4.3.1   Platform: ProCon—a Prototype of a Social Discussion Platform for Controversial Topics

We built a prototype named *ProCon*. It is a social discussion platform where people contribute comments on controversial topics, which other users could browse. Upon logging into the system, users would first see a front page with

Table 4.1: **Research questions and findings in Study 2 (Chapter 4)**

	Research question	**Findings**
RQ1	How do fine-grained position indicators impact users' selective exposure (**RQ1a**)? How does accuracy motivation mediate the impact (**RQ1b**)? How is the effect related to using fine-grained instead of dichotomous positions (**RQ1c**)?	Position indicators decreases selective exposure, but only for people with high accuracy motivation. The main reason is that the fine-grained position indicators facilitate those with high accuracy motivation to attend to moderately different opinions.
RQ2	How do position indicators impact perception of the position discrepancy (**RQ2a**), and agreement (**RQ2b**)?	Position indicators help users differentiate moderately different opinions from the more extreme ones, and reduce their disagreement with the former.
RQ3	How doe position indicatorS impact users' knowledge gain on either side?	By reducing selective exposure of users with high accuracy motivation, position indicators facilitate them to gain more knowledge regarding the opposite position.

a list of controversial topics, from which they could choose to start browsing (order of the topics is randomized). After clicking on a topic, users would enter the main page of the comment list, as shown in Figure 4.2 (top).

For each topic, it presents a short description on the controversy and a list of comment snippets with author names. The snippet provides a summary of the author's opinions on the topic. Users can click the snippet and a comment page will be opened on a new tab (4.2 bottom). Users can click "read more comments from the user" to continue reading more comments from the same author. On the comment page, users are asked to rate the commenter's knowledge and position, and their agreement with each comment.

4.3.2 Material

I selected 6 out of the 8 topics identified in the first study (see Appendix A). Since I recruited participants from the similar population, these topics were likely to have a balanced distribution of opinions. Two undergradu-

Figure 4.2: **ProCon interface (baseline) used in Study 2 and Study 3 (Chapter 4 and Chapter 5): comment list (top) and comment page (bottom)**

ate assistants and I worked on collecting and labeling the material. To be ecologically valid, we obtained the material from real user posts on the Internet. We started by collecting 150-200 comments, with 60-100 words each, expressing varied opinions on each of the topics, from online forums including debate.org, procon.org and Yahoo! Answers[1].

We then worked on categorizing these comments into four categories: strong pro, moderate pro, moderate con and strong con (see Table 6.2 for examples). The inter-rater reliability (Fleiss Kappa) reached 0.85, which is considered good agreement [153]. Inconsistently labeled comments were either discussed to reach consensus, or excluded. We categorized the comments with the following features to be of moderate positions: 1) Expressing indecisiveness by mentioning merits of both sides, but leaning towards one side; 2) Supporting one side but with certain condition (e.g., "it is good only if certain restriction applies"); 3) Uncertain tone (e.g., "I guess", "maybe"). In contrast, strong comments expressed one-sided and affirmative opinions, and were often with confident and strong tone (e.g., "I believe", "definitely").

I created fictional commenters (i.e., author of the comments) for the experiment to control for the consistency of information and distribution of positions. For each topic, I created 32 commenters, with 8 under each of the four categories: strong pro, moderate pro, moderate con and strong con. For each commenter, I selected 3 comments from the corresponding comments pool to be shown on the comment page (they will be randomly loaded one at a time when participants click on "read more from the commenter"). As a priority, I grouped comments collected from the same real commenter together. Otherwise, I carefully selected and modified the comments, if necessary, to make them reasonably consistent for a particular commenter. For each commenter, I created an "opinion summary" (see Table 6.2) to be shown as the snippet on the comment list interface. The order of authors and the order of comments are randomized for each participant.

---

[1]www.debate.org, www.procon.org, answers.yahoo.com

Table 4.2: **Examples of comment for the topic "should prescription drug be allowed to directly advertise to consumers?"**

Category	Comment summary	Comment
Strong pro	Drug advertisements inform patients about medical issues therefore should be encouraged.	An important benefit of direct to consumer advertising is that it fosters an informed conversation about health, disease and treatments between patients and their health care practitioners. Pharmacy members want patients and consumers to talk to their physicians about the medicines that may help them and to fully understand the known risks regarding these medicines.
Moderate pro	I do not have problem with advertisement about prescription drug as long as it provides unbiased information.	I do not have problem with advertisement about prescription drug as long as it provides unbiased information, meaning including both its benefits and side effects in a honest manner. A prescription drug is something that consumers should be making a rational decision about. And the more information consumers have, the better decisions they make.
Moderate con	We should ban, or at least limit the advertisement of prescription drugs to avoid people making biased medical decisions.	I think that U.S. should limit the television commercials for prescription medications for this may influence the doctors and make a medication "more popular" without proper reason. Excessive promoting of medications using commercials may influence the "popularity" of a certain medication, then people would more frequently ask for it.
Strong con	Drug advertisement is dangerous and should be banned as it may mislead patients.	Direct to consumer prescription drug ads, like most advertisements, are intended to sell the product being advertised. Such ads use marketing tactics that manipulate, create false impressions, and mislead consumers instead of educating them about the drugs. It can be dangerous if patients start self-diagnosis by the information they get from advertisement.

Figure 4.3: **ProCon interface with position indicators**

4.3.3 Experimental Design and Procedure

To test the effect of position indicators, I conducted an experiment with repeated-measure between-subjects design. That is, half of the participants were randomly assigned to the baseline condition using the interface shown in Figure 4.2. The other half used an interface with the fine-grained position indicators as shown in Figure 4.3— a "position bar" that uses lengths of different colors to represent both the valence and the magnitude of position reflected in the messages. Each participant completed 6 tasks using the system.

Upon arriving at the lab, participants were given a questionnaire to collect their demographic information and topic relevant information, which will be discussed in detail in the "measurement" section. They were then briefed about the experiment platform and the task—to freely browse the comment list and contribute their own comments at the end of each task. Participants then logged into the comment list and completed the 6 topics in any order they preferred. For each topic, after finishing browsing the interface, they proceeded to the page where they wrote down their comments and finished answering topic relevant questions. It took 40-80 minutes for participants to finish all the experiment tasks.

4.3.4 Participants

By posting recruiting ads in email newsletters, I recruited 32 participants in the Urbana-Champaign area in Illinois, the United States. 20 of them are undergraduate or graduate students. The rest are a mix of faculties, university staffs, engineers, etc. All the participants were randomly assigned to the group with position indicators (group 1) and the control group (group 2). No significant difference in gender (group 1 43.8% male, group 2 37.5% male), age (group 1: M=32.0, SD=13.3, group 2: M=24.6, SD=8.5; $p = 0.28$), education (25% in group 1 and 31.3% in group 2 are graduate students or have graduate degree, the rest are undergraduate or have bachelor degree) or political leaning (scale 1-conservative to 5-liberal, group 1: M=3.5, SD=0.9, group 2: M=3.5, SD=1.1; $p = 0.85$) was observed between the two groups.

4.3.5 Measurements

In this section, I introduce variables that were measured directly by pre- and post- task questionnaires and the experiment platform. Specific indexes used for each analysis will be discussed in the result section. In the pre-experiment questionnaire, attitude, accuracy motivation, and topical knowledge were measured for each topic. After each task, on the page where participants submitted their comments, their topical attitude and knowledge were measured again. The experiment platform was able to log all the interactions for analyses of information selection.

Attitude Index

The same 5-item semantic differential scale used in the first experiment (see Appendix B) was used to measure the participants' attitudes on each controversial topic [127]. I calculated the average ratings of the five items to be the participant's *prior attitude index* for the topic (Cronbach's $\alpha = 0.96$, considered good internal consistency).

Accuracy Motivation Index

In this experiment, I directly measured the participants' topic related accuracy motivation by two items:

1) *How much are you interested in learning more about the topic?*

2) *How much do you desire to know the truth about the topic?*

The ratings were based on a 1-none to 7-a lot scale, and were averaged to calculate the accuracy motivation index (Cronbach's $\alpha = 0.88$, considered good internal consistency).

Topical Knowledge

For each topic, I used a recall task to measure the participants' general knowledge about each topic. I asked each participant to write down "what reasons or arguments immediately come to your mind for people who support or oppose the issue" using simple sentences such as "religious reason". They were asked to write pro-arguments and con-arguments in two separate bulleted lists. The same question was asked again after the browsing task. I counted the number of points that appeared in the post-experiment but not in the pre-experiment questionnaire as a proxy of the participant's *knowledge gain* after browsing the website.

4.4   Results

Participants in the condition with position indicator selected a mean of 5.1 (SD=2.7) commenters and read 1.8 comments (SD=0.8) from each commenter for each topic. Participants in the control condition selected a mean of 5.3 (SD=2.7) commenters and read 1.7 (SD=0.8) comments from each commenter. No significant difference in the number of commenters ($p = 0.69$) or comments per user ($p = 0.60$) was observed between the two groups.

To start with, I coded each participant's prior attitude on each topic to be pro or con based on whether his or her prior attitude index was below or above 4, which was the neutral point of the scale and also the median among all the participants. 20 out of the 192 cases where the participants scored exactly 4 were removed from the analysis. Then I coded the positions of each commenter shown on the interface to be extremely consistent, moderately consistent, moderately inconsistent or extremely inconsistent based on whether their opinions were on the same or opposite side of the participant's prior attitude, and their attitude magnitude.

In the rest of the section, I first analyze whether position indicators could nudge to reduce selective exposure, and how users' accuracy motivation mediated the effect. I also further examine the selection of each type of information to understand whether using fine-grained position indicators, i.e., showing the magnitude of position, made a difference in the selectivity (RQ1). Note that I examined the selection at both the commenter and comments level, and the conclusions were similar. So in this section I focus on reporting the results for the commenter selection, and briefly mention the latter. I then look at how the position indicators impacted the perceived position of information (RQ2). Unlike the first experiment, I was not able to find significant results in attitude change, potentially because the pre- and post- study attitude questionnaires were given closely in time. Instead, I chose to focus on examining the knowledge gain, especially knowledge about the opposing side (RQ3), as an outcome measure of exposure to diverse perspectives.

4.4.1 Information Selection (RQ1)

To answer RQ1, how position indicators impacted the participants' selective exposure tendency, I created a selective exposure index (SEI), calculated by the difference in the number of attitude consistent and attitude inconsistent commenters selected. A positive value indicated the exhibition of selective exposure, and the magnitude of the index reflected the size of the difference.

Position Indicators on Selective Exposure (RQ1a)

I started by examining the indicators' overall impact on the participants' selective exposure. I ran a mixed-effect regression on the selective exposure index by having the presence position bars (present=1, absent=0) as the fixed-effect independent variable (N=172), and participants as random effects. I found a main effect of the presence of the position bars ($\beta = -0.51, t(30) = -2.20, p = 0.04$), suggesting that, overall, position bars decreased the participants' selective exposure.

Accuracy Motivation Mediating the Effects of Position Indicators (RQ1b)

In RQ1b, I ask whether position indicators have differential impact on users with varied levels of accuracy motivation. To answer this question, I included the accuracy motivation index in the mixed-effect regression model on selective exposure index. Interestingly, I found a significant interaction between accuracy motivation and the presence of position bars on the selective exposure index ($\beta = -0.37, t(28) = -2.22, p = 0.03$). It answered the research question in the affirmative— the accuracy motivation mediated the effect of position indicators on decreasing selective exposure.

In Figure 4.4, I visualize the interaction between accuracy motivation and the presence of position bars, by first preforming median splits to create the high or low accuracy motivation groups, and presenting the mean SEI for each group.[2] The visualization highlights that: The presence of position bars decreased the selective exposure for topics in which participants had higher than median accuracy motivation. However, the position bars had little impact for participants who had lower than median accuracy motivation. The results suggested that *the position bar was effective in mitigating selective exposure, but only for those who were motivated to accurately learn about the topic.*
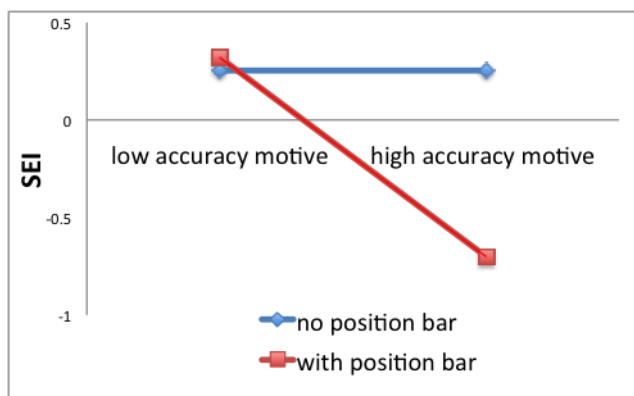
Figure 4.4: **Average Selective Exposure Index (SEI) for topics with high/low accuracy motivation, with/without position indicators**

To confirm this effect, I also looked at whether the same pattern of re-

---

[2]The median splits were only performed for the visualization. The accuracy motivation index was treated as a continuous variable in all regression models.

sult was observed in the selection of the individual comments, since participants had the option to read from 1 to 3 comments from each commenter. I calculated the selective exposure index based on the number of comments read. I observed the same trend that there was a significant interaction between accuracy motivation and the presence of position bars ($\beta = -0.68, t(28) = -2.20, p = 0.03$), suggesting that the above-mentioned conclusion held.

Valence and Magnitude of Position on Information Selectivity (RQ1c)

I examine the information selectivity in more detail to understand whether using fine-grained position categories, to provide both the valence and the magnitude of the information stance, benefited the participants' information seeking process. I ran a mixed-effect regression on the percentage of the type of commenters selected for each topic, by each participant (N=668, 4 categories for each of the 172 tasks), by including information position valence (consistent=1, inconsistent=0), position magnitude (moderate=0, extreme=1), the participant's accuracy motivation index, and presence of position bars (present=1, absent=0) as fixed-effect independent variables [3]. I found a significant three-way interaction among the presence of position bars, accuracy motivation, and the position valence of information ($\beta = -0.08, t(16) = -2.95, p < 0.01$); and a significant three-way interaction among the position bars, accuracy motivation, and the position magnitude of information ($\beta = -0.06, t(16) = -2.07, p = 0.04$). These significant interactions suggest that both the valence and magnitude reflected in the position indicators affected the participants' information selectivity, and the effect differed for participants with varied levels of accuracy motivation.

To unpack these interactive effects, given that I was interested in how the position bars would affect selective exposure, I separated the selected commenters into two groups: attitude consistent commenters and inconsistent commenters. I then ran the regression analysis for each group to study how the presence of position bars affected the selection pattern of users with extreme or moderate positions (i.e., position magnitude), for participants with varied levels of accuracy motivation.

---

[3]participants were always included as a random effect

For attitude inconsistent commenters, I ran a mixed-effect regression model on the selection percentages by including the presence of position bars, the accuracy motivation index, and the information position magnitudes as independent variables. I found a significant three-way interaction among all the independent variables ($\beta = -0.06, t(24) = -2.18, p = 0.03$). Figure 4 provides a clearer picture of the interaction— *the position bar facilitated participants who had high accuracy motivation to explore arguments supporting both the moderate and extreme attitude-inconsistent stances, but participants with low accuracy motivation were less interested in those with the moderate stance.*

To verify the conclusion, I ran a mixed-effect regression on the selection percentages for moderately inconsistent commenters only, by including the accuracy motivation index and the presence of position bar as independent variables. Results showed a significant two-way interaction between the two ($\beta = 0.05, t(28) = 2.72, p < 0.01$). No such effect was shown for selecting extremely inconsistent opinions. The result suggests that showing the position magnitude, i.e., using a fine-grained position indicator, was what made the difference in selective exposure between users with high and low accuracy motivation.
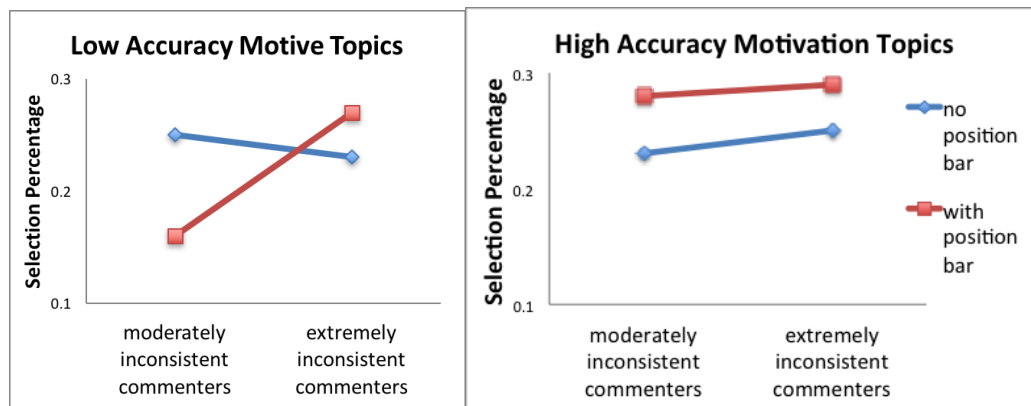
Figure 4.5: **Average selection percentages of attitude-inconsistent commenters with moderate versus extreme stances, by participants with high/low accuracy motivation**

Meanwhile, by analyzing the total number of commenters selected, I found that accuracy motivation had a marginally significant positive effect ($\beta = 0.22, t(30) = 1.69, p = 0.08$), showing that participants selected fewer com-

menters when they had low motivation to accurately learn about the topic. Given the low learning motivation, it is likely that they were only willing to spend limited effort exploring the topic, and therefore might not be interested in knowing and weighing the detailed arguments, hence tended to avoid the information with moderate stances. In contrast, the greater motivation to form an accurate understanding of the topic might have encouraged participants to seek detailed arguments from commenters of varied positions, both the moderate and extreme ones.

For attitude consistent opinions, I ran the same mixed-effect regression model but did not find the same three-way interaction to be significant. I found a significant two-way interaction between presence of position bars and accuracy motivation ($\beta = -0.03, t(27) = -1.93, p = 0.05$)— this simply means that they interacted in decreasing selective exposure, the same conclusion as illustrated with the SEI. I also found a main effect of information position magnitudes ($\beta = 0.09, t(27) = 4.52, p < 0.01$), showing that, unlike with attitude-inconsistent information, regardless of the accuracy motivation level, participants consistently preferred commenters with extremely consistent over those with moderately consistent positions.

At the comment level, I performed the same analysis with the percentages of comments as the dependent variable. I found a significant four-way interaction among presence of position bars, accuracy motivation, information position valence, and magnitude ($\beta = 0.08, t(16) = 1.93, p = 0.05$). Closer examination revealed similar patterns as with the selection of commenters: With the position bars, participants with high accuracy motivation read more moderately inconsistent comments than participants with low accuracy motives, which resulted in the former group's overall lower selective exposure tendency. This difference in attending to comments of moderate stances was not observed for attitude consistent comments.

In summary, the above results show that using a fine-grained position indicator had benefits in reducing selective exposure. The most important result was that *showing the position magnitude of information could "nudge" people with high accuracy motivation to explore moderately inconsistent opinions. As a result, the position indicators increased the exposure to attitude-challenging information for participants with high accuracy motivation, but had no such impact for participants who had low accuracy motivation.*

4.4.2 Information Judgment (RQ2)

After examining the effect of the position bars on the participants' selective exposure, the next question is to what extent the position bars could lead to changes in their reception of attitude inconsistent information. For reception, I examined two information judgments—position and agreement, both to reflect the perceived discrepancy between the information and one's own position, with larger discrepancy indicating more unfavored reception.

Position Judgment (RQ2a)

To answer RQ2a—how do the aspect indicators influence the position judgment, I examined the participants' position ratings (on a 1-pro to 5-con scale) given to each commenter after they read the comments. First, I looked at the correlation between the participants' judgments and the "ground truth" of the position labels (coded as extremely pro commenter=1, moderately pro commenter=2, moderately con commenter=3, extremely con commenter=4). I found that the position bars increased this correlation from 0.64 to 0.78 ($Z = 4.77, p < 0.01$). It suggested that after seeing the position bars, the participants' judgments of the stances of the commenters were closer to those indicated by the position bars.

Relative Position Index

In the hope of improving users' reception of attitude-inconsistent information, my main focus was to examine relative position judgment–judgment of an information position as relative to one's own position. To do so, I first calculated the average position judgment ratings given to the four categories of commenters (extremely/moderately consistent or inconsistent). Since the position judgment used a 1-pro to 5-con scale, if the participant had positive prior attitude, the *Relative Position Index* given to each type of commenter would be equal to the average position judgement rating. If the participant had negative prior attitude, the *Relative Position Index* would be calculated by 6 minus the average position judgment rating. The Relative position index reflected how discrepant the participant perceived the position of the information is from his or her side, with a higher number indicating a higher discrepancy.

Results of Relative Position Judgment

I ran a mixed-effect regression model on the relative position index (N=535, we removed the 153 cases where the corresponding type of commenters were not selected) by including the presence of position bars, information position valences and magnitudes as independent variables. I found a significant three-way interaction among all the independent variables ($\beta = -0.78, t(24) = -2.46, p = 0.01$). All the other two-way and main effects, except for that between presence of position bars and commenters' position valences, were also significant. I unpacked the three-way interaction by running a mixed-effect regression on the relative position index among attitude consistent commenters and inconsistent commenters separately. I found a significant two-way interaction between the opinion magnitudes and the presence of position bars in the attitude inconsistent group ($\beta = 0.59, t(28) = 3.10, p < 0.01$), but not in the attitude consistent group ($\beta = -0.21, t(28) = -0.98, p = 0.33$). In other words, the three-way interaction was caused by the fact that the position bars led to significant differences in the position judgments of commenters with attitude-inconsistent positions, but not of those with consistent positions.

Figure 4.6 illustrates this result: With the help of position bars, participants were better at differentiating between information with extremely and moderately attitude-inconsistent stances. Indeed, by examining moderately inconsistent commenters only, the main effect of the presence of position bars on the relative position index was significant($\beta = -0.33, t(28) = -2.00, p = 0.05$), suggesting that the position bar made participants to judge this group of commenters (moderately inconsistent) to be closer to themselves, compared to those who did not see the position bar. This is consistent with the notion that *people are inclined to perceive arguments supporting an opposite position to be more extreme than they actually are. The fine-grained position bar seemed to be able to mitigate such bias.*

Agreement (RQ2b)

To further study how the position indicators influenced the reception of diverse opinions, RQ2b asked whether the position bars influenced the par-
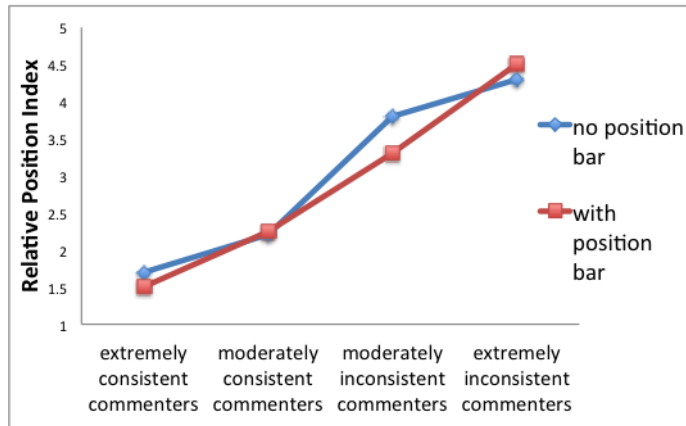
Figure 4.6: **Average Relative Position Index given to information supporting varying positions**

ticipants' agreement with information supporting the varying positions. To answer this question, I investigated the participants' agreement ratings given to the comments (by giving an agreement rating under each comment). I expected that agreement ratings, compared to the position judgment, would be impacted more by the participants' own position magnitude. For example, participants with more extreme prior attitudes would likely have stronger disagreement with commenters who took the opposite positions. Also, if the participant was only moderately leaning towards one side, he or she might agree more with opinions that were also in the moderate positions. Therefore, I examined the participants' agreement ratings by taking their prior attitude extremity into consideration.

Agreement Index and Prior Attitude Extremity Index

The *Agreement Index* was calculated by the average agreement ratings the participant gave to each type of comments—extremely consistent, moderately consistent, extremely inconsistent, and moderately inconsistent.

The *Prior Attitude Extremity Index* was defined as the distance of one's prior attitude index from the neutral point, and was calculated by the absolute value of the prior attitude index minus the neutral value 4. For example, if a participant scored 5 in the prior attitude index, he or she would be coded as con prior attitude with prior attitude extremity index of 1.

Results of Agreement

66

I ran a mixed-effect regression model on the agreement index (N=535), with the presence of position bar, prior attitude extremity, information position valence and magnitude as independent variables. I found a significant three-way interaction among presence of position bars, prior attitude extremity index, and information position valence ($\beta = -0.86, t(16) = -3.65, p < 0.01$), and a significant three-way interaction among presence of position bars, prior attitude extremity, and information position magnitude ($\beta = -0.47, t(16) = -2.12, p = 0.04$). The result suggests that position bars differentially influenced users with moderate and extreme prior attitudes on their agreement with the comments of different positions.

To unpack the interactive effects, I separated the consistent and inconsistent commenters into two groups. For attitude inconsistent commenters, I ran a mixed-effect regression model on the agreement index by including presence of position bars, prior attitude extremity index and information position magnitude as independent variables, and found a significant three-way interaction among them ($\beta = -0.41, t(24) = -1.91, p = 0.05$). Figure 4.7 [4]explained the three-way interaction we found above: The group of participants with extreme prior attitudes tended to show stronger disagreement with attitude-challenging information than those with moderate prior attitudes. However, the fine-grained position bar was able to help them identify comments that are only moderately different and improve their agreement with them.

To further verify this conclusion, I looked at the agreement index for moderately inconsistent comments only. Indeed, I found the two-way interaction between prior attitude extremity and presence of position bars to be significant ($\beta = 0.65, t(28) = 3.73, p < 0.01$). For the agreement index for extremely inconsistent commenters, I only observed significant main effect of prior attitude extremity ($\beta = -0.46, t(31) = -5.89, p < 0.01$), as participants with extreme prior attitudes, as compared to the group taking more moderate positions, tended to disagree more strongly with the extremely inconsistent comments, regardless of the presence of position bars.

For attitude consistent commenters, I ran the same mixed-effect regres-

---

[4]Same as previous figures, median splits on prior attitude extremity were used to illustrate the interaction in the figure, but included as continuous variable in the regression models.

sion model and did not observe the above-mentioned three-way interaction, but found a significant two-way interaction between the participants' prior attitude extremity and information position magnitudes ($\beta = 0.31, t(28) = 2.89, p < 0.01$). This was consistent with the expectation that, regardless of the presence of position bars, participants tended to agree more with information that is similar to their position. That is, those with more extreme prior attitudes tended to agree more with confirmatory information taking more extreme stances, while those who had moderate prior attitudes tended to agree more with information supporting similarly moderate positions.
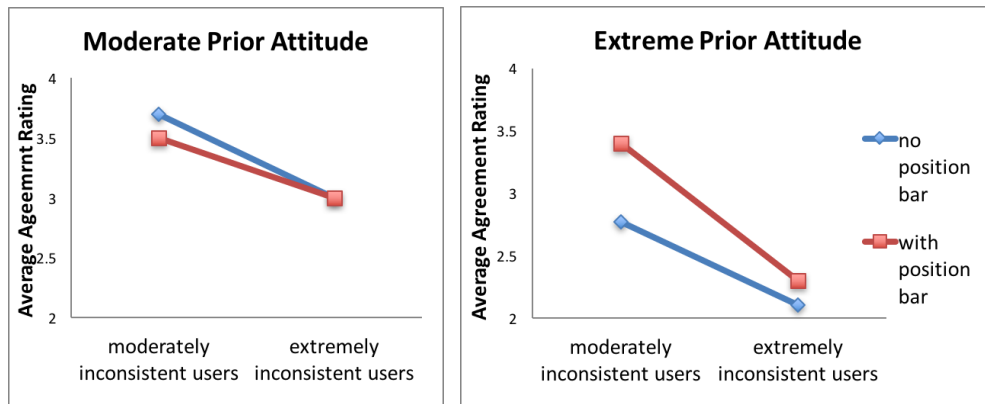
Figure 4.7: **Average Agreement Index for attitude-inconsistent information by participants with moderate/extreme prior attitudes**

The above analyses performed on position judgment and agreement revealed a consistent theme: The fine-grained position indicators were helpful for participants to differentiate attitude-challenging information with only moderate stance from the extreme one. With that, the indicator moderated the participants' tendency to over-estimate the discrepancy between the position of attitude-challenging information and their own positions. Taken together, the results suggest that *showing the position magnitude of information not only helped reducing selective exposure, but also improved the reception of attitude-challenging information that was only moderately different.*

4.4.3 Knowledge Gain (RQ3)

An important goal of exposing people to diverse viewpoints is to promote awareness of arguments from different perspectives. Last but not least, I analyzed the participants' knowledge gained on both sides as the outcome measurement of exposure to diverse opinions. In particular, I was interested in whether the difference in selective exposure, as a result of introducing position indicators as a central nudge, led to differences in knowledge gained about both sides. Therefore, I focused on testing the interactive effect between accuracy motivation and the presence of position bars.

In the survey before the experiment, I asked each participant to list arguments that could "immediately come to your mind" that people make to support or oppose the topic. After the experiment, they were asked to answer the same question again. I compared the two answers from each participant and counted how many attitude-consistent arguments and inconsistent arguments, respectively, appeared in the post-experiment questionnaire but not in the pre-experiment one. I used this number as a proxy measure of knowledge gain from using the system.

I ran a mixed-effect regression model on the knowledge gain index for arguments supporting the opposite side of the participants' prior attitude (attitude inconsistent knowledge gain) by including presence of position bars and accuracy motivation index as independent variables. Consistent with the result of selective exposure, I found a significant interaction between the two ($\beta = 0.33, t(28) = 3.10, p < 0.01$). Figure 4.8 illustrated the two-way interaction with median splits of accuracy motivation index: With the help of position bars, participants gained more knowledge about the opposite side for topics they had high accuracy motivation, but not for those they had low accuracy motivation. For knowledge gained for arguments supporting their prior position, I did not observe any significant effect of position bars or accuracy motivation.

The result is consistent with the finding that the position bar nudged participants with high accuracy motivation to attend to more attitude inconsistent information. The results therefore further supported the conclusion that fine-grained position indicators was an effective "central nudge" that could reduce selective exposure, and support deliberative learning, for people who
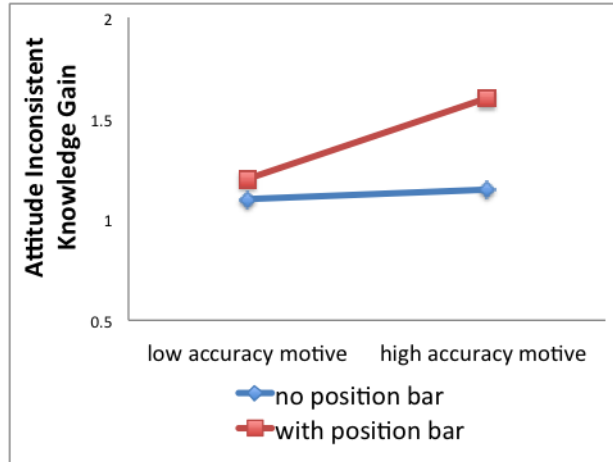
Figure 4.8: **Average Knowledge Gain about arguments supporting the opposite position**

were motivated to accurately learn about the topic.

4.5 Discussion

The study makes two main contributions. First, it demonstrates that interface features providing not only the valence but also the magnitude of information position are useful for encouraging exposure to diverse viewpoints. In particular, the results show that distinguishing information that is only "moderately different" could increase exposure and improve reception of information supporting the competing views, and as a result, lead to better learning outcomes, especially better awareness of diverse perspectives. While I studied the effect of a simple position bar, my goal is to highlight the value of a set of interface features that categorize information into more fine-grained positions rather than a dichotomous model. Potential relevant design examples include OpinionSpace's [21] visualization interface that map the stances of information to spatial distance, the slider bar used by Consider.It [29] for users to actively explore arguments made by people of varied positions, and the system that is similar to the one used by DataPotrait [102] to explicitly recommend information that leans towards different positions but shares certain common grounds.

Second, the results provide validation for the ELM-based design guideline I proposed. It demonstrates that central nudge, i.e., information cue reflecting content characteristics, is more effective in reducing selective exposure for those with a higher level of accuracy motivation. While some recent research pointed out the co-existence of diversity-seeking and challenge-averse users, this study suggests that there may be underlying mechanisms that drive such behavioral differences—the motivation to accurately understand the topic. For users with high accuracy motivation, there may be inherent user needs for designs that better facilitate the deliberative choices of information and careful exploration of the opinion space. Future research should further explore identifying the moderating variables of selective exposure, and how they mediate the effectiveness of nudging designs.

This work also aligns with taking a more fine-grained view of attitude-relevant information by identifying what kind of attitude-inconsistent information is more desirable, especially for serving the accuracy motivation. And here I suggest a potential candidate is information that is less extreme in its position, considering and acknowledging both sides, and providing distinct insights that bridge diverse perspectives. Such insight is not only useful for designing information presentation, but can also inform the design of information retrieval systems. To serve the diversity-enhancing goal, these systems should look beyond attitude relevance as the primary user preference, by identifying information that is both challenging and appealing.

From a technical perspective, two challenges emerged in this set of conclusions: How to infer users' accuracy motivation levels; and how to infer the position magnitude of information. Learning from research on personalization technologies, users' accuracy motivation can be potentially identified from both explicit inquiry and implicit inference based on their user profiles or previous behaviors [154]. There are also conditions where people are generally more likely to be motivated to accurately learn about a topic. For example, decision-making support tools will more likely target users who have high accuracy motivation, compared to, e.g., hedonic use of social media. It is also possible for designers to actively increase the users' accuracy motivation by, for example, emphasizing the social norm of valuing diversity [104] or highlighting the utility of diverse or challenging information [51].

To infer information stances, various techniques have been developed to

classify positions (e.g., ideological leaning, consumer opinions) based on either information content (e.g., machine learning, opinion mining, sentiment analysis [155])) or non-text features, such as user voting or social network features [71, 156]. One future direction would be to develop information retrieval methods that can accurately identify information with "moderate positions", and arguments that bridge different perspectives.

Chapter 5

STUDYING EXPERTISE INDICATORS AS A PERIPHERAL NUDGE

People's information seeking process is seldom rational. Instead, we are frequently using "shortcuts" to decide where to look for information we want. These mental shortcuts include checking where the information comes from, what other people say about it, and how it is presented. Psychology and communication scholars call them *peripheral features* of information, as they are not an integral part of the information content, and the dual-process theories suggest that they tend to be processed in the peripheral route.

In the Internet age, peripheral features of information have become more prevalent than ever. For example, we check Facebook post that receives many "likes". We follow people with many other followers on Twitter. We look for articles shared by authorities and friends. We tend to trust information from a website that is professionally designed. Relying on these heuristics help us navigate through the overloaded online information. On the one hand, the digital media affords the modalities and spaces to design various peripheral features. On the other hand, contemporary social technologies make it possible to incorporate "social cues" such as likes, sharing, etc. to become useful peripheral futures that are not only used by information seekers, but also information retrieval algorithms. These social peripheral features often fall into three (not necessarily exclusive) categories: 1) Popularity cues, which show large number of people attended to the information (e.g.,number of views); 2) Endorsement cues, which show other people acknowledged the value of the information (e.g., "Liked" or shared by a friend or by many); and 3) source competence cues, which signal the merits of the information source (e.g., voted experts).

An abundance of studies manifested that these peripheral features are able to alter users' information-seeking paths, whether or not these studies explicitly target information behavior changes. For example, Salganik et al.

demonstrated that users' choices of items on a music downloading website were largely influenced by "social cues" that showed how many times others have downloaded the song. Knobloch-Westerwick[157] showed that "collaborative filtering" cues, such as average rating by others and times viewed by others, could significantly impact users' selectivity on news websites. Kim and Sundar [147] showed that popularity cues on a health message board not only nudged people to attend to the more popular information, and also led them to see the information more favorably. Research on social navigation [158] explored more sophisticated ways to present various forms of social endorsement that could shape and improve users' information seeking behaviors [159, 160].

So theoretically, we should be able to increase users' tendency to attend to attitude-inconsistent information by adding positive peripheral nudges to the information. Several recent studies provided evidence that social endorsement cues, including the number of "likes"[105], annotation indicating shared by similar people [106], increase users' interest in reading partisan news that they may disagree with. However, I argue that this may not be enough to reduce selective exposure. Ecologically speaking, peripheral nudges are likely to happen for all sides of views. While peripheral features such as popularity cues may nudge to increase users' interest in reading attitude-inconsistent information that is popular, they would also increase interest in reading attitude-consistent information that is popular. Although it is possible to customize these cues based on individuals' positions, peripheral features are usually not selectively presented to only attitude-challenging information on the common platforms.

The key question that should be asked, instead, is *what kind of peripheral nudge can nudge more for attitude-inconsistent than for attitude-consistent information*, and thus reducing the selective exposure to (at least for certain type of) information? In other words, the effectiveness of the peripheral nudge should interact with the information position. I note that, in Messing and Westwood's study [105], there was some evidence in the data showing that the nudging effect of social endorsement cues is stronger for disagreeable information, and thus it is possible to reduce selective exposure among news with higher social endorsement. In Munson and Agapie's study, they did not observe any interactive effect between news agreeableness and social cues

[106].

What about other types of peripheral nudge? Psychological research suggests that source competence (credibility, expertise, etc.) related cues are a promising choice in creating differential effect for information with varied positions. Aronson et al.[150] found an interactive effect between source credibility and position discrepancy (distance between recipient's and message's position): The preference for high credibility over low credibility sources became more prominent with increasing position discrepancy. Other studies [161, 162] also repeatedly found that, when people are consistently predisposed with the message's advocacy, a source with low expertise is not necessarily less favorable than a source with high expertise; however, when inconsistently predisposed, people consistently prefer, or are only persuaded by, sources with high expertise.

5.1 Overview of the Study

Inspired by this line of evidence on the interaction between preference for source competence and position consistency, I propose to study interface cues that indicate source expertise as a type of peripheral nudge to change users' information behavior towards exposure to high-quality diverse views. I also hypothesize that, if there is indeed an interactive effect between peripheral nudge and information position on information selectivity, it is possible that such effect will be reinforced by providing an explicit label of the information stance. This is what was done in Messing and Westwood's study [105], where the differential nudging effect of social endorsement for agreeable and disagreeable news was reported when both social endorsement and information position cues were present. Based on the idea, I conducted a 2×2 experiment to study how the expertise indicators interacted with position labels to affect users' information behaviors, including both information selection and information reception. Furthermore, since expertise indicators are considered to be a type of peripheral nudge by enlisting the source expertise related heuristics, I explored whether accuracy motivation mediated its effect. Specifically, I asked the following research questions in Table 5.1:

Table 5.1: **Research questions and findings in Study 3 (Chapter 5)**

	Research question	**Findings**
RQ1	How do expertise indicators, by themselves and together with position indicators, influence selective exposure?	Together with position indicators, expertise indicators reduced selective exposure to information from expert sources, but increased it for that of non-expert sources. The reason is that high expertise indicators are stronger nudges for attitude-inconsistent than for consistent information.
RQ2	How do the expertise indicators influence the users' agreement with information of varied positions (**RQ2a**) and perceived source expertise (**RQ2b**)?	Expertise indicator increase the agreement with, and the perceived source expertise of information, even though information did not actually differ in quality or expertise.
RQ3	How do accuracy motivations mediate the effect of expertise indicator on reducing selective exposure?	No significant mediating effect of accuracy motivation is found for the nudging effect of expertise indicators.

5.2 Methodology

This study used the same platform and procedure as the experiment discussed in Study 2 (Chapter 4). So I will not repeat the sections of platform, material, procedure, and measurement described in the last chapter. I will discuss experimental design and participants of this experiment in this section.

5.2.1 Experimental Design

The experiment adopted a 2×2 repeated-measure experimental design with the presence or absence of position bars, and that of expertise indicators as independent variables. That is, participants were randomly assigned to one of the four conditions: 1) with neither position nor expertise indicators; 2) with only position indicators; 3) with only expertise indicators; and 4) with both position and expertise indicators. Figure 5.1 illustrates the interfaces used in condition 3 and condition 4. Condition 1 is the baseline interface shown in Figure 4.2, and condition 2 is the interface with aspect indicators

shown in Figure 4.3.

I used a "star badge" design as expertise indicators (see Figure 5.1). I introduced to participants that the stars were calculated by the collective votes of other users of the ProCon platform, as everyone could rate the commenter expertise when they browse the comments page. The same indicators were shown on the comment page. Note that since position magnitude was not the focus of this study, I would only consider pro versus con positions, and consider extreme and moderate ones as the same position category.
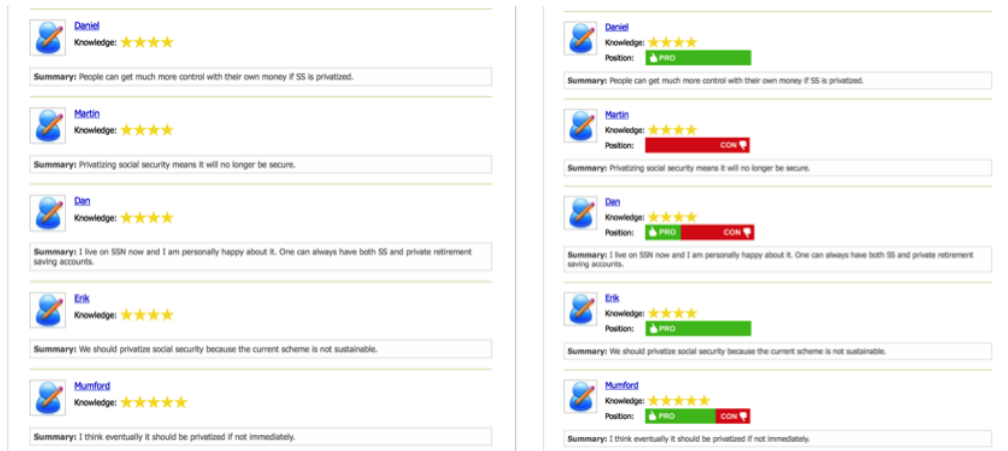
(a) ProCon with expertise indicators (condition 3)

(b) ProCon with both expertise and aspect indicators (condition 4)

Figure 5.1: **ProCon interfaces with expertise indicators**

I randomly assigned and evenly distributed the 32 commenters to have four types of expertise indicators: 1-star, 2-star, 4-star and 5-star. I did not manipulate the actual content quality of their comments. The reason was that, to answer RQ2 and RQ3, I was interested in the extent to which a simple interface cue could impact the participants' judgment of information. Therefore I controlled for the content quality. In reality, there is also no guarantee that the information quality and the source competence cues will always match. For example, reputation systems commonly used in online forums often do not perfectly reflect the quality of the content produced, as regular users who post frequently tend to be ranked higher than experts who post only occasionally.

77

5.2.2   Participants

76 participants were recruited from a college town in Illinois, USA. They were randomly assigned to the four experiment conditions. On average, participants aged 24.20 (SD=9.74). 55.3% were female. 78.9% were undergraduate or graduate students. The rest consisted of professors, university staff, engineers, etc. On a scale of 1-conservative to 5-liberal, the participants' political leaning scores averaged 3.55 (SD=0.92). Data from one participant in condition 1 (control condition) was removed due to technical errors during recording.

5.3   Results

On average, participants selected 5.18 (SD=2.54) commenters for each topic. No statistically significant difference in the number of selected commenters was seen between the four conditions. To analyze the participants' information selectivity with regard to attitude consistency, I categorized the commenters to be attitude consistent or inconsistent based on the participant's prior attitude. This was done in the same fashion as in the previous experiments. I first coded each participant's prior attitude to be pro or con based on whether his or her prior attitude index was above or below 4. Then I coded each commenter's attitude consistency according to whether the commenter was on the same or the opposite side of the participant's prior attitude.

To analyze the expertise-based selectivity, I categorized all commenters with 4 or 5 stars to be experts, and those with 1 or 2 stars to be no-experts (to be precise, they were only "indicated" to be experts or non-experts. I will call them "experts" and "non-experts" throughout the chapter for simplicity reason),. Thus, among the 32 commenters, there were eight under each of the four categories: attitude-consistent experts, consistent non-experts, inconsistent experts and inconsistent non-experts.

In this section, I will first report the results on selective exposure (RQ1), then examine in detail how expertise indicators impacted the selectivity of attitude consistent and attitude inconsistent information differently, to understand the changes in selective exposure. I will then analyze how expertise

indicators impacted the participants' information reception, including their agreement with (RQ2a), and perceived source expertise of (RQ2b), information with varied positions. Lastly, I explore whether accuracy motivation mediates the effect of expertise indicators on changing the selective exposure tendency (RQ3).

5.3.1 Selective Exposure (RQ1)

RQ1 asks how expertise indicators, by themselves and together with position bars, impacted the participants' selective exposure tendency. To this end, I first examined the effects of the presence of expertise indicators and/or position bars on the overall selective exposure, then analyzed it among expert commenters and non-expert commenters separately.

Overall Selective Exposure

To start with, I calculated the overall selective exposure index for each topic, by each participant, defined as the difference between the number of attitude consistent and inconsistent commenters selected. A positive value of the index indicated the presence of selective exposure, and the magnitude reflected the extent.

To examine the effects of position and expertise indicators on the participants' selective exposure, I ran a mixed-effect regression model on the selective exposure index by including the presence of expertise indicators (present=1, absent=0) and the presence of the position bar (present=1, absent=0) as the fixed-effect independent variables. As with the previous chapters, participants were always included as a random-effect variable in all regression models. I found a significant negative main effect of the presence of position bar ($\beta = -0.60, SE = 0.16, t(72) = -3.74, p < 0.01$; Average SEI of conditions with aspect indicators is -0.20, conditions without aspect indicators is 0.40), suggesting that, consistent with conclusions in Study 2 (Chapter 4), the multi-level position indicators decreased the participants' selective exposure tendency.

No statistically significant effect of the presence of expertise indicators

was observed ($\beta = 0.03, SE = 0.16, t(72) = 0.21, p = 0.83$), suggesting that expertise indicators did not affect the overall selective exposure. In the next section, I will examine the selective exposure to information indicated to be from experts and non-experts separately.

Selective Exposure for Experts and Non-Experts

To first check whether participants exhibited different selective exposure tendency for information from expert and non-expert sources, I started by comparing the selective exposure index for expert and non-expert commenters. To do so, I calculated two selective exposure indexes for each topic, for each participant: One is calculated by the difference between the number of attitude consistent experts and inconsistent experts selected (expert selective exposure); and the other in the same way for the non-experts (non-expert selective exposure). I ran a mixed-effect regression model on the expert/non-expert selective exposure index by including the presence of expertise indicators, that of position bars, and the type of commenters (expert=0, non-expert=1), as fixed-effect independent variables. I found a significant three-way interaction among all the independent variables ($\beta = -1.05, t(67) = -2.98, p < 0.01$), suggesting that the selective exposure differed between selecting expert commenters and non-expert commenters when both expertise and position indicators were presented.

Figure 5.2 illustrates the results by plotting the average SEI for each condition. *While neither the expertise indicator nor the position indicator alone led to a significant difference in selective exposure, having both indicators significantly decreased selective exposure to information from expert sources, but increased it for non-expert sources.*

To verify the conclusion, I tested the effect of expertise and position indicators on the expert and non-expert selective exposure index separately. Indeed, I found a significant negative interaction between the presence of expertise indicators and position indicators ($\beta = -0.58, t(71) = -2.04, p = 0.04$) on expert selective exposure index, and a significant positive interaction between the two on non-expert selective exposure index ($\beta = 0.47, t(71) = 1.98, p = 0.05$).
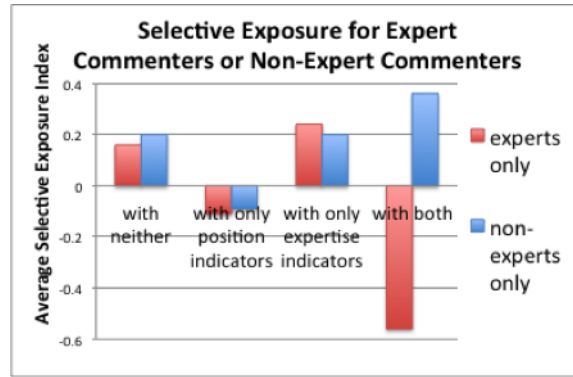
Figure 5.2: **Average expert and non-expert Selective Exposure Index (SEI) in each condition**

In summary, I found that by providing source expertise cues, and meanwhile explicitly labeling the position of the information content, the participants' selective exposure was decreased for information from expert sources, but increased for non-expert sources. It is an encouraging result suggesting that source expertise indicators could serve as a positive nudge for seeking high-quality, yet attitude-challenging information. However, the effect seemed to be only present when the content position is explicitly shown. This suggests that the nudging effect of expertise indicators may interact with content position, which was reinforced by the presence of explicit position labels. In the following section, I examine in detail how expertise indicators impacted the selectivity of attitude-consistent and inconsistent information differently.

Expertise-Based Selectivity

To begin with, I checked whether there was a statistically significant difference between attitude-consistent and attitude-inconsistent information in how the expertise indicators influenced the expertise-based selectivity, i.e., selection of between information from expert v.s. non-expert sources. I ran a mixed-effect regression model on the number of each type of commenters selected, by including the commenters' expertise (1-expert, 0-non-expert), attitude consistency (1-consistent, 0-inconsistent), presence of expertise indicators (1-present, 0- absent), and position bar (1-present, 0-absent) as the

fixed-effect independent variables. Table 5.2 shows the statistically significant terms in the model.

Table 5.2: **Significant terms in the regression model for expertise-based selectivity**

	β	t-value	p-value
Intercept	1.17	8.04	$< 0.01^{**}$
Expertise indicators	-0.59	-2.90	$< 0.01^{**}$
Expertise $\times$ Expertise indicators	1.25	6.42	$< 0.01^{**}$
Attitude consistency $\times$ Expertise $\times$ Expertise indicators $\times$ Position indicators	-0.94	-2.45	0.01^{*}

The positive two-way interaction between expertise and presence of expertise indicators suggested that expertise indicators nudged users to select more information from expert sources. The four-way interaction among all independent variables implied that the nudging effect differed for attitude-consistent and attitude-inconsistent information, with the presence of position bar. To further unpack the interactive effect, I looked at the expertise selectivity for attitude-consistent and attitude-inconsistent information separately.

Expertise Selectivity for Attitude Consistent Information

Looking at attitude consistent information only, I ran a mixed-effect regression model on the number of commenters selected, by including the commenters' expertise, the presence of position bar and expertise indicators as fixed-effect independent variables. I found a significant three-way interaction among all the three independent variables ($\beta = -0.81, t(71) = -2.93, p < 0.01$). The two-way interaction between expertise and the presence of expertise indicators was still significant ($\beta = 1.29, t(71) = 6.57, p < 0.01$).

Figure 5.3a shows the average number of attitude consistent experts and non-experts selected in each condition. It reveals an interesting observation that explains the three-way interactions mentioned above: *Expertise indicators nudged participants to select more information from expert sources. However, when adding position bar, this tendency became weaker for selecting attitude consistent information.* To verify the conclusion, I tested the interactive effect of the commenters' expertise and the presence of posi-

tion bars in the conditions with expertise indicators (condition 3 and 4). As expected, I found a significant negative interaction between the two ($\beta = -0.70, t(36) = -3.43, p < 0.01$), and a positive main effect of the commenters' expertise ($\beta = 1.21, t(36) = 8.35 p < 0.01$).

Expertise Selectivity for Attitude Inconsistent Information

Among attitude inconsistent information, I ran the same mixed-effect regression on the number of commenters selected, by including commenters' expertise, the presence of expertise indicators and that of position indicators as fixed-effect independent variables. I did not observe the same three-way interaction ($\beta = 0.13, t(71) = 0.48, p = 0.63$), but only a significant two-way interaction between expertise and the presence of expertise indicators ($\beta = 1.24, t(71) = 6.42, p < 0.01$). These results suggested that, as illustrated in Figure 5.3b, participants consistently selected more information from experts than non-experts for attitude inconsistent commenters. Adding position indicators did not significantly changing, even slightly increased this tendency.
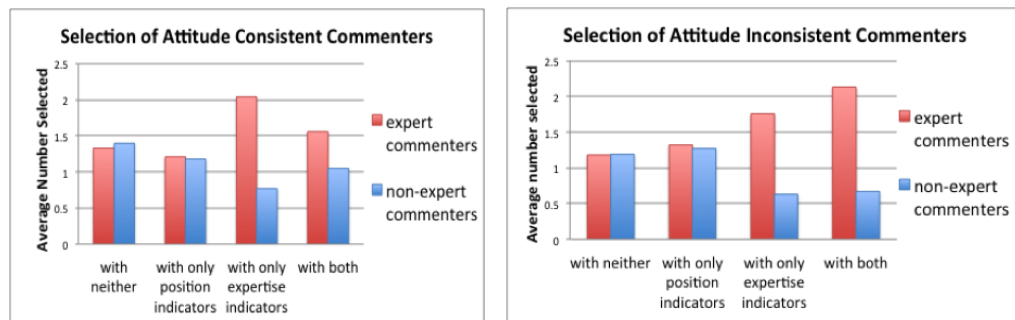
(a) Selection of experts v.s. non-experts for attitude inconsistent commenters

(b) Selection of experts v.s. non-experts for attitude consistent commenters

Figure 5.3: **Expertise-based selectivity for attitude consistent versus inconsistent information**

To summarize, the above results showed that high expertise indicators nudged users to select more information, and the nudging effect interacted with the information position, when the position was explicitly labeled. Specifically, the nudging effect was weaker for attitude-consistent than for inconsistent information. This result echoed the conclusion from the literature [150, 161, 162], that information seekers' preference for sources of high

competence over low competence is more pronounced for selecting dissonant information than consonant information. Note that the results also suggested that such varied preference for source expertise may only become pronounced when the content position is explicitly labeled, so that users did not have to examine the snippets to interpret the opinion. An alternative explanation is that participants might generally see both high expertise indicators and consistent position indicators as positive cues. In the conditions with only expertise indicators, the participants' selectivity was dominated by expertise cues, but in the condition with both expertise and position indicators, the participants' selectivity was influenced by both. Thus, expertise cue might had lower impact on selectivity.

The interactive effect between the presence of expertise indicators and the information position explained the results on selective exposure observed above. *When both expertise indicators and position bars were presented, high expertise indicators nudged more for seeking attitude-inconsistent information than attitude-consistent information, and thus reduced the selective exposure to information from expert sources, and meanwhile, increased it for information from non-expert sources.*

5.3.2   Information Judgment (RQ2)

While expertise indicators may induce heuristics at the information selection stage, the participants' judgment of information should be based on the contents after they read them. RQ2 asks whether expertise indicators still had impact at the post-reading stage on the participants' agreement with, and perceived expertise, of information.

Agreement (RQ2a)

Participants were asked to rate their agreement with each comment they read. I created the *agreement index* by calculating the average agreement ratings the participant gave to each type of comment—from attitude-consistent experts, consistent non-experts, inconsistent experts and inconsistent non-experts. I ran a mixed-effect regression model on the agreement index

(N=1170 [1]) by including attitude consistency, expertise, and the presence of expertise indicators as fixed-effect independent variables. Table 4 shows the statistically significant results.

Table 5.3: **Significant terms in the regression model for agreement ratings**

	β	t-value	p-value
Intercept	3.08	39.77	$< 0.01^{**}$
Attitude consistency	0.67	12.81	$< 0.01^{**}$
Expertise $\times$ Expertise indicators	0.29	2.72	$< 0.01^{**}$

As illustrated in Figure 5.4, the main effect of attitude consistency means that participants always had higher agreement with attitude-consistent commenters. The two-way interaction between indicated expertise and presence of expertise indicators showed that high expertise indicators significantly increased the participants' agreement with the information.

Figure 5.4: **Average Agreement Index for different types of information in conditions with or without expertise indicators**

To verify, I tested the main effect of the presence of expertise indicators on the agreement with information with high or low expertise indicators separately, and found its positive effect for information with high expertise indicators only ($\beta = 0.21, SE = 0.09, t(73) = 2.21, p = 0.03$). Moreover, the absence of interaction between attitude consistency, expertise and presence of expertise indicators ($\beta < 0.01, SD = 0.21, t(73) = 0.02, p = 0.99$)

---

[1]426 cases where participants didn't select the type of commenters (therefore gave no ratings) were removed from analysis

implied that, high expertise indicators evenly increased agreement with both attitude-consistent and attitude-inconsistent information.

In summary, I found that *high expertise indicators nudged people to agree more with the information, even though the information content did not actually differ in quality or source expertise. Such positive effect on agreement happened consistently for both attitude-consistent and attitude-inconsistent information, although people always agreed more with information that aligns with their positions.*

Expertise Judgment (RQ2b)

While the agreement rating reflected a user's evaluation of information from a subjective point of view, when asked to evaluate others' expertise, one should be expected to adopt a more objective view to evaluate the information source. RQ2b asked to what extent the expertise indicators impacted the perceived expertise after participants read the contents. Meanwhile, it would also be interesting to see whether position discrepancy led to the derogation of source expertise.

In the experiment, when exiting the comment page, participants were asked to rate the commenters' topical expertise. I calculated an *expertise judgment index* by the average expertise ratings given to each type of commenters by each participant. I ran a mixed-effect regression model on the expertise judgment index by indicating the attitude consistency, indicated expertise, and the presence of expertise indicators as fixed-effect independent variables (N=1170). Table 5 shows the statistically significant results.

Table 5.4: **Significant terms in the regression model for expertise ratings**

	β	t-value	p-value
Intercept	3.28	39.79	< 0.01**
Attitude consistency	0.24	4.61	< 0.01**
Expertise $\times$ Expertise indicators	0.70	6.54	< 0.01**

As illustrated in Figure 5.5, The main effect of attitude consistency suggested that participants tended to give higher expertise ratings to attitude

consistent sources than to attitude inconsistent ones. The two-way interaction between expertise and the presence of expertise indicators suggested that the high (low) expertise indicators increased (decreased) the participants' perceived expertise of the information sources, even though the commenters were randomly assigned to be experts or non-experts in the experiment.
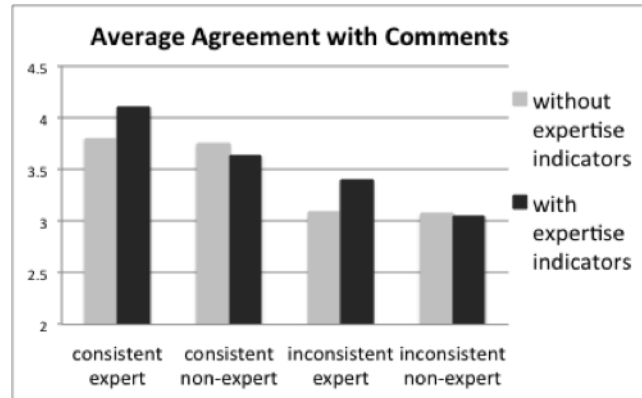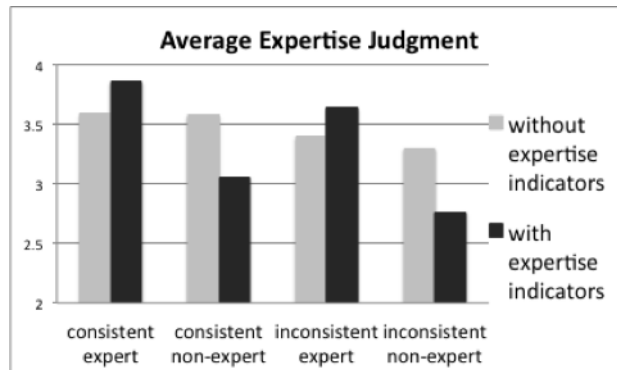

Figure 5.5: **Average Expertise Index for different types of information in conditions with or without expertise indicators**

To verify, I tested the main effect of the presence of expertise indicators on the expertise judgment for information indicated to be from experts and non-experts separately. As expected, I found its positive main effect for expert commenters ($\beta = 0.24, t(73) = 2.31, p = 0.02$), and negative main effect for that of non-expert commenters ($\beta = -0.44, t(73) = -3.28, p < 0.01$). The absence of interaction between attitude consistency, expertise, and presence of expertise indicators ($\beta = 0.06, t(73) = 0.28, p = 0.78$) suggested that the expertise indicators evenly impacted perceived expertise for attitude-consistent and inconsistent information. It also pointed out that, even when explicitly indicated to have equal level of expertise, attitude consistent commenters still tended to be viewed as more knowledgeable than attitude inconsistent ones. To verify this conclusion, I tested the main effect of attitude consistency on the perceived expertise for expert commenters and non-expert commenters separately, in the conditions with expertise indicators (condition 3 and 4). Indeed, I found its positive main effect on perceived expertise for both expert commenters ($\beta = 0.22, t(37) = 2.80, p < 0.01$) and non-expert commenters ($\beta = 0.26, t(36) = 2.00, p = 0.05$).

In summary, I found that expertise indicators affected the participants' expertise judgment. Specifically, high expertise indicators increased their per-

ceived source expertise of information, and low expertise indicators decreased that, even though the contents of the two did not differ in quality or expertise. However, participants systematically considered attitude-consistent information to have higher source expertise than attitude inconsistent one, even when the two were indicated explicitly to have equal level of expertise.

5.3.3 Mediating Effect of Accuracy Motivation (RQ3)

In this section, I explore whether accuracy motivation mediated the nudging effect of expertise indicators on reducing selective exposure. According to the analyses above, the most encouraging nudging effect of expertise indicators is that, together with position indicators, it would reduce selective exposure to information from expert sources (Figure 5.2 experts only). This is specifically illustrated by the positive interactive effect between the presence of position indicators and expertise indicators on the Selective Exposure Index (SEI) for information from expert sources. Therefore, to explore the mediating effect of accuracy motivation, I included the participants' accuracy motivation index in this particular regression model. The result showed no significant mediating effect of accuracy motivation, with the absence of three-way interaction between the presence of expertise indicators, position basr, and accuracy motivation index (=-0.07, SE=0.18, t(67)=-0.38, p=0.71).

In Figure 5.6, I show the descriptive results of average SEI for information from expert sources in all the four conditions, separately for participants with high or low accuracy motivation (by performing median split on accuracy motivation index). It demonstrates that *expertise indicators, together with position bars, successfully nudged to seek more attitude-inconsistent information from expert sources for both participants with high and low accuracy motivation.*

5.4 Discussion

Peripheral features such as source competence and popularity cues can be used as "nudges" to change the users' information behavior in desirable di-
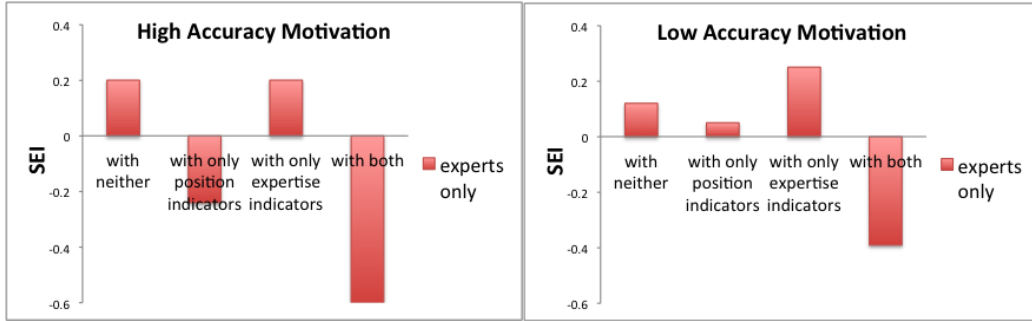
Figure 5.6: **SEI to information from expert sources of participants with high versus low accuracy motivation**

rections. To the best of my knowledge, this is the first study to show that peripheral nudges can also interact with content characteristics, i.e., create stronger nudging effect for certain type of information than others, and thus shaping information behaviors in more complex ways. Specifically, I show that the nudging effect of expertise indicators interacts with information position. This is consistent with conclusions from psychological research demonstrating that source competence related heuristics would impact information selectivity more for dissonant than for consonant information. Leveraging such tendency, the expertise indicator was able to reduce selective exposure to information from expert sources.

I also show that a simple feature suggesting high expertise could have a lasting effect on the users' information reception, by increasing their perceived expertise of, and agreement with, the information. The result encourages incorporating source expertise rating systems (e.g., driven by content, user profile or social feedback) in information and social technologies that can accurately reflect the source competence and information quality. Not only may it aid the users' information seeking process to efficiently focus on the most informative messages, even if they only want to attend to a limited amount of dissonant information, but also, it may potentially improve the reception of dissonant information and encourage people to seek common grounds with divergent views.

Moreover, I explored the mediating effect of accuracy motivation and found that the nudging effect of expertise indicators did not significantly differ for people with high and low accuracy motivation. This is encouraging, suggest-

ing that peripheral nudges that induce positive heuristics can work to expose people to more high-quality attitude-inconsistent information, even for those with low level of accuracy motivation, and are often prone to selective exposure.

Future work can explore other peripheral features as peripheral nudges, such as popularity cues and endorsement cues. It is worth pointing out that, in this study I particularly leveraged the interactive effect between source expertise and content position to attain the nudging effect on reducing selective exposure for expert sources. Other peripheral features may or may not have such interaction effects. Still, we may explore other ways to leverage peripheral nudges to shape the users' information selectivity in more desirable ways. For example, we may selectively customize peripheral nudges to target only attitude-inconsistent information.

Given the power of peripheral nudges, information retrieval techniques should consider them as an integral part of information. Obviously the idea is not new, as we frequently see systems that prioritize items that are popular, endorsed by others or from competent sources. But to better utilize them to nudge the users' information seeking paths, we should use them more deliberately, by considering their interactive or combined effect with characteristics of the information content. For example, we may design algorithms that favor less attended information with positive peripheral cues to diversify the users' information diet.

From a different angle, given the power of peripheral nudges for changing both the information selection and information reception, we should design them with caution. Nowadays, as by-products of social features, many peripheral nudges are not created intentionally, nor well-understood for its potential effect. We need to be cautious about whether positive nudges are consistently given to desirable information, whatever the desirability means. Most time the answer is probably positive, because these nudges are often, by nature, associated with heuristics that are supposed to aid the information seeking process. For example, endorsement, popularity, and source competence cues usually reflect some desirable features of the information such as usefulness, quality, etc.

However, sometimes the matter is more complicated. One potential prob-

lem is that common social peripheral nudges are often products of "homogeneous social activities", as endorsed by people who hold similar views, by "liking", sharing, etc. The outcome is that, in realistic settings, peripheral nudges are often not given equally to information of varied positions as in the above controlled experiment. So peripheral nudges can be biased. If information that expresses certain views tends to come with stronger peripheral nudges—for example, because people who hold such views tend to use the social features more actively, they will likely increase its reachability and favorability. Such implication of inequality for different viewpoints is detrimental to the goal of promoting exposure to diverse perspectives. I will empirically explore the problem in Study 5 (Chapter 7), by studying how different opinion groups may use social features differently, and thus creating unequal peripheral nudges for their viewpoints.

Chapter 6

EXPLORING SELECTIVE EXPOSURE IN CONSUMER HEALTH INFORMATION SEEKING AND ASPECT INDICATORS AS A CENTRAL NUDGE

Previous empirical research on selective exposure used tasks that largely varied in domains and contexts. For example, many looked at information gathering after participants made decisions such as purchasing choices, policy preferences, candidate selections, etc.[42, 79]; some looked at people's preferen with regard to their own ideological leaning [11, 60]; and some looked at patients' preference and avoidance of health information [163]. With the observation, scholars pointed out that there is a lack of understanding of how task domains impact people's selective exposure, which may be the key to reconciling the varied empirical results on its existence and extent [164, 165]

Meanwhile, the accuracy-motivation framework may provide some theoretical ground to explain how selective exposure varies in different contexts. For example, consistent with the prediction of reduced selective exposure with lower level of defense motivation, research found that in hedonic contexts, information selectively is less impacted by attitude consistency [166]; and that when people are only observer[126] or advisor [167] , instead of the decision-maker, they tend to seek more balanced information. Conversely, one may predict more pronounced selective exposure in threat-prone task contexts, as both my previous experiment (Study 1 in Chapter 3) and other studies (e.g., [49]) have demonstrated that threat contributes to defense motivation, hence increases selective exposure tendency.

One potential example of a threat-prone domain is health-related information seeking, where considerations of suffering, risk, even mortality are often involved. Indeed, as I have discussed in the literature review section, many have identified selective exposure to be a problem that may impair medical decisions [107], health risky behavior changes [111], and engagement in unbiased information seeking, especially for patients with high-risk, life-threatening diseases such as cancer [108, 109, 110], or when mortality is in

the picture [167].

However, the situation may become more complex when considering accuracy motivation. Given the outcome importance regarding one's health status, in theory, people should have high accuracy motivation in many health-related contexts. For this reason, unlike everyday news consumption, an interesting potential attribute of health information seeking is that high outcome importance often co-exists with high threats. It is unclear whether people would engage in more pronounced selective exposure when seeking medical information online, especially when facing high-risk decisions. It thus merits empirical examination, particularly to understand, when high-risk health decisions are likely to incur both high threats and high accuracy motivation, which one will prevail.

In Study 2 (Chapter 4), I demonstrated that accuracy motivation mediates the effectiveness of central nudges. An interesting question would be whether such conclusion still holds in the health information seeking context. To explore, I study another type of central nudge in this chapter—"aspect indicators", which show the issue aspect each piece of information focuses on. I also explore whether disease risk mediates its impact on information selectivity. The aspect indicator is a type of a central nudge because it strictly reflects characteristics of the information content, and can provide support for making deliberative choices of information. The expectation that it may serve as a nudge for seeking attitude-challenging information is based on the observation that people holding different attitudes often focus on different aspects of a controversial issue. So by differentiating information that focuses on different aspects, it may separate information regarding "non-conflicting" aspects where users may not have biased beliefs, and thus potentially reduce users' resistance to it, even if it advocates a different position.

Aspect indicators may be especially helpful for making multi-attribute decisions [168], where the decision outcome depends on the assessment and integration of different attributes of the decision. Labeling the varied aspects may encourage one to explore, and weigh the pros and cons of different attributes (i.e., aspects), which can be an important step in making informed decision and "debiasing" people with biased beliefs [130]. Medical decision is considered a typical multi-attribute decision. It requires the decision maker to, first of all, be aware of the various aspects of the candiadate choices, e.g.,

93

effectiveness, side effects, availability, price, etc. Then the decision maker has to be able to gather accurate and balanced information for each attribute. Moreover, how one weighs varied aspects in the decision may depend on the circumstances. For instance, a previous work conducted by Bergus et al. [169] suggests that, for serious diseases, people may prioritize the effectiveness of the medicine. While for common, mild diseases where there are many treatment options, side effects may also become a key consideration.

As an exploration of selective exposure in the health domain, I studied the problem in the context of social opinion seeking by using a comment aggregator (similar to Yelp) that supports comparative medical decisions. While decades of research have been done on clinical medical decision support systems (DSS), studying patients oriented DSS is a fairly new topic. According to Hibbard et al.[170] , consumer medical decisions often have the following characteristics: 1) Challenges in using information containing technical terms and complex ideas; 2) Need to compare multiple options on several variables; 3) Require to weigh the various factors according to individual values, preferences and needs. To support such activities, Hibbard suggests tools that consumer medical decision-making support should be designed with the goal of lowering cognitive effort, and one way to do so is to present decision makers with other patients' real-life experience, which is usually easier to comprehend than formal medical information. Consistent with the idea, many proposed that the large volume of medical related contents on social media, including drug review websites, online health communities, etc. can be a useful information resource for consumer health information needs [100, 171]. This study represents an effort in this direction.

6.1   Overview of the Study

In this experiment, I studied the effects of aspect indicators in reducing selective exposure in opinion seeking for medical decisions, particularly for aspects where one did not have conflicting beliefs. I chose to focus on the two most critical aspects of medical decisions— effectiveness and side effects, by studying the situation where users had biased beliefs regarding only one of the two aspects (*biased aspect*). I compared the results for tasks regarding

high-risk diseases and those of low-risk diseases (*disease risk*), to explore the role of disease risk as moderators of selective exposure in the medical context. Furthermore, as the last one of the serial studies based on the ELM-based design guideline, by considering the aspect indicator as a central nudge, I examined the mediating role of disease risk. Specifically, I asked the following research questions listed in Table 6.1:

Table 6.1: **Research questions and findings in Study 4 (Chapter 6)**

	Research question	**Findings**
RQ1	Is there selective exposure in seeking medical opinions? How does disease risk and biased aspect impact the selective exposure?	Although selective exposure is generally low, it becomes pronounced in one situation—having biased beliefs in the effectiveness of treating high-risk diseases. The analysis suggests that the criticality of biased aspect matters, as effectiveness is the more critical aspect for treating high-risk diseases. The result presents a warning that selective exposure may become more pronounced in more threatening situations in consumer health information seeking.
RQ2	How do aspect indicators influence selective exposure? How does disease risk mediates the effect?	Aspect indicators decrease selective exposure in the "unbiased aspect" for high-risk diseases, but have little effect in tasks for low-risk diseases, suggesting that as a central nudge it is has stronger impact on selective exposure in high-risk medical contexts, where accuracy motivation tends to be high. However, I also found evidence that the impact may be further mediated by the criticality of aspect, since it is less effective when the unbiased aspect happens to be the critical aspect of decision.
RQ3	How does selective exposure impact decision bias? How do aspect indicators moderate the impact?	The situation with the most pronounced selective exposure—concerning about the effectiveness for high-risk disease—results in the highest decision bias. Aspect indicators are able to moderate the problem by reducing selective exposure in the "unbiased aspect".

6.2 Methodology

In this section, I will introduce the third information aggregator prototype we developed for studying medical opinion seeking. I will then discuss the procedure, experimental design, content material, measurements and participants in this study.

6.2.1 Platform: CompareMed—a Prototype for Drug Review Aggregator

I created CompareMed, a prototype of comment aggregator system that intended to support comparative medical decisions by presenting peer patients' comments. The system was introduced as an information aggregator that, when a user searched for two treatments he or she wanted to compare, it aggregated relevant comments from health-related social media. The system presented comments for each treatment side by side. For each treatment, it presented a list of comment snippets (with usernames), which one could click to read the whole comment. Each snippet was shown with a star rating, which reflected how negative or positive the comment was evaluating the treatment, using a 1 (negative) to 5 (positive) scale. A screenshot of the interface is shown in Figure 6.1

6.2.2 Procedure

The experiment consisted of 6 comparative medical decisions, each regarding a particular disease. Before the experiment, participants were given a short survey to collect their demographic information. They were then briefed about the CompareMed system. Upon logging into the system, they would see a list of the six disease tasks, which they could choose to start in any order they preferred.

Unlike in the political information-seeking tasks, participants did not come with existing preferences and beliefs about the two treatments. To introduce "pre-existing biased beliefs" in the experiment, I adopted a priming approach. For each disease task, participants were given a scenario to imag-
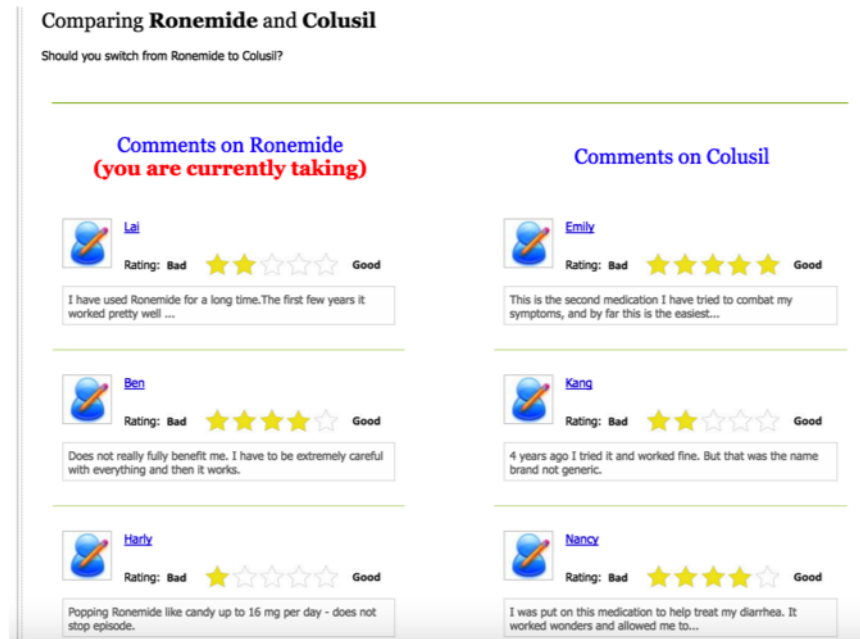
Figure 6.1: **CompareMed Interface**

ine that a close friend was having negative experience with a treatment he or she just started, and facing the decision on whether to switch to an alternative treatment, which the friend heard might work better. To reinforce the bias against the current treatment, the task description is accompanied with a picture depicting the relevant negative experience (e.g., showing someone suffering from stomach pain for the disease "diarrhea"). Participants were asked to use CompareMed to learn about the two options and provide suggestions to the friend. I expected that it would introduce *biased beliefs favoring one treatment (alternative) to the other (current)*.

Participants were asked to freely browse the system until they were ready to give suggestions. Then they proceeded to the survey page to rate their preference between the two treatments and wrote down suggestions for the friend. The system automatically logged the interaction data for analyses.

6.2.3 Experimental Design

The experiment included two between-subjects variables: biased aspect (effectiveness/ side effects), the presence of aspect indicators (presence/ ab-

sent); and one within-subject variable: disease risk (high/ low). First of all, I studied the effect of aspect indicators. Half of the participants were randomly assigned to use the baseline experiment platform shown in Figure 6.1. The other half used the one with aspect indicators— star signs that not only showed the comment sentiment, but also, by the labels ("side effects" and "NO side effects", or "ineffective" and "effective"), indicated whether the comments focused on the effectiveness or side effects of the treatment. Different color schemes were used to further differentiate the two types of comments. In contrast, in the control condition, only one kind of star sign was used, with generic labels ("bad" and "good")(Figure 6.2). This format, including the star signs and separate ratings to evaluate different aspects of a medication, is commonly seen on drug-review websites such as WebMD, RxList, etc.
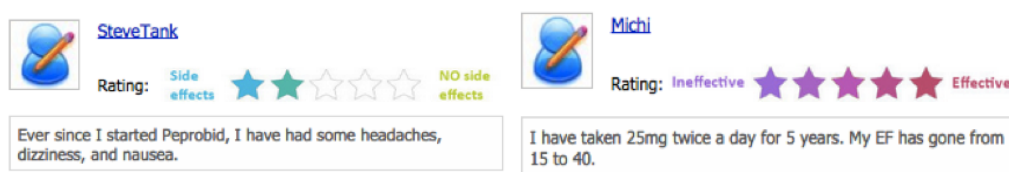
Figure 6.2: **Snippets of side effects comment (left) and effectiveness comment (right) with aspect indicators**

My hypothesis was that aspect indicators could reduce selective exposure for *aspects where one does not have biased beliefs*(I will call it "unbiased belief" for simplicity). To test the hypothesis, it required that participants would only have pre-existing bias on only one of the aspects. I therefore included "biased aspect" as a between-subjects variable. That is, participants were randomly assigned to one of the four conditions: biased beliefs on effectiveness/side effect, and with/without aspect indicators. To manipulate biased beliefs on a particular aspect, I designed two versions of the task scenario. Half of the participants would read that the friend was trying the current treatment but did not see it relieved the symptoms, and heard that the alternative option might work better to cure the disease. The other half would read that the friend was suffering from certain side effects, which I ensured to be consistent side effects mentioned in the comments, and heard the alternative option would not cause similar problems. I also used different pictures that consistently depicted either the symptoms or the side effects

described in the scenario.

To compare the results for high versus low risk diseases, disease risk was deigned to be a within-subjects variable. Each participant completed six tasks. Half of them were about high-risk diseases that could be potentially life-threatening and needed to be cautiously controlled: *congestive heart failure (CHF)*, *deep venous thrombosis (DVT)*, and *acute asthma attack*. The other three were low-risk, minor diseases that could often be treated or controlled by over-the-counter medicines: *diarrhea*, *back pain* and *heartburn (acid reflux)*. To reinforce the perceived risk level, for high-risk diseases, a short description of the disease and its potential risk was included in the task description. No such information was given for low-risk diseases.

6.2.4 Material

For each disease, I chose two medicines that received similar user ratings on WebMD as candidate treatments to be compared in the experiment. Their names and other identifying information were revised to avoid recognition.

For each medicine, an undergraduate assistant and I collected 50-60 comments from the "user comments" sections of popular drug information websites (e.g., WebMD.com, rxlist.com), and medical discussion forums (e.g., medhelp.org). We were cautious about having a balanced number of positive and negative comments, and comments regarding either effectiveness or side effects. We intentionally excluded comments that did not have a clear focus on either aspect.

We independently rated the comments for their sentiment (positive/ negative) and focused aspects (effectiveness/ side effects). We excluded comments that we disagreed on and ended up with 32 comments for each medicine, with evenly distributed sentiment and aspects in order to create a "controversial" decision situation where the information did not significantly favor one option over the other. Table 6.2 presents examples of each type of comment.

To avoid order effect, in each experiment session, one of the two medications was randomly chosen to be the current option, and consistently described in the task scenario. On the screen, the current and alternative option were randomly placed on the left or right side, with the current option

Table 6.2: **Examples of comments presented by CompareMed**

Aspect	Position	Comment
Effectiveness	**Positive**	Dosage currently 25mg of Peprobid twice a day and I am feeling pretty great. My EF is up to 50-55% from 20%. And my BP is 130/85, which is a great improvement from 208/182. My heart doctor and my PCP both told me that this will be a miracle medicine and they were right
	Negative	Diagnosed with Congestive Heart Failure, It was hoped this would reduce my BP and help me avoid worse conditions up ahead. Now I have been taking Peprobid for the last 3 months, I really can not say it has made me feel one bit better.
Side effects	**Positive**	After being on Atenolol for 5 years, switching to Peprobid was a vast improvement. I used to jitter and feel exhausted all the time with the other meds. No side effects so far with Peprobid.
	Negative	I am only on 12.5 mg of Peprobid but I sleep all night & day, if I am not sleeping I am so exhausted I can not move with out becoming so short of breath I have to stop and sit not moving for at least 15-29 min. The side effects are dragging me down.

labeled as "currently taking" on the heading (see Figure 6.1). The order of the 6 disease tasks and the order of comments were also randomized.

6.2.5 Measurement

In this experiment, I focused on analyzing information selection and decision outcome. Information selection was logged by the system. To measure the decision outcomes, after each browsing session participants were asked to rate two statements: 1) The alternative is a better option; and 2) I would like to suggest to switch to the alternative; based on a 1 (disagree) to 5 (agree) Likert scale. The ratings were averaged to create a *decision index*, which reflects the participants' preference for the alternative over the current option.

In addition to the common demographic information such as age, gender, and education, I measured the participants' experience with seeking health information online by asking them to rate how often they look online for (based on a 1-never to 5-very often scale): 1) information about medicines; 2) disease related information; 3) healthy lifestyle related information; and 4) health-related social media. I also asked them to self-rate their knowledge for each of the disease used in the experiment (based on a 1-know nothing to 5-know a lot scale).

6.2.6 Participants

I recruited 67 participants from the Urbana-Champaign community by using online job boards and a campus-wide mailing list. Participants' average age was 26.58 (SD=11.58), including 21 male and 13 with post-graduate degree. There was no significant difference in age, gender, education, self-rated disease knowledge, and experience of health information seeking online for participants assigned to the four experimental groups.

6.3 Results

In the experiment, participants selected an average of 11.00 (SD=6.22) comments for each task. To start with, I coded attitude consistency for each comment. Given the introduced bias of "an unfavorable current treatment and a potentially better alternative", I coded *negative comments on the current* and *positive on the alternative* treatment to be *attitude consistent* (note in each task the two treatment was randomly chosen to be the current or alternative treatment), and coded *positive comments on the current* and *negative on the alternative* to be *attitude inconsistent*. Therefore, for each task, there were four types of comments: attitude consistent/inconsistent comments on effectiveness/side effects.

Manipulation Check

Before analyzing the result by considering biased aspect—the aspect where one was primed with biased beliefs—as an independent variable, I conducted a manipulation check by testing whether the selectivity between comments on effectiveness and side effects differed for those receiving different scenario. The intuition was that if participants were successfully primed to be concerned about one aspect but not the other, they would be selecting more comments regarding the former.

I started by calculating the selection percentage of effectiveness/side effects related comments (divided by the total number of selected comments) for each task, by each participant. I ran a mixed effect linear regression model on the selection percentage, by including comment aspect (effectiveness/side effects), the participant's biased aspect (effectiveness/side effects), and the presence of aspect indicators (present/absent) as fixed-effect independent variables. For all analyses in this chapter, participants were included as a random-effect variable. I found a significant three-way interaction among all the fixed-effect independent variables ($\beta = 0.13, t(64) = 2.43, p = 0.02$). I illustrate the interactive effect by plotting the average selection percentages of effectiveness and side effects related comments in Figure 6.3.

The figure shows that, as expected, with the presence of aspect indicators, participants selected more comments for the aspect on which they were primed with biased beliefs. To confirm, I tested the interactive effect between
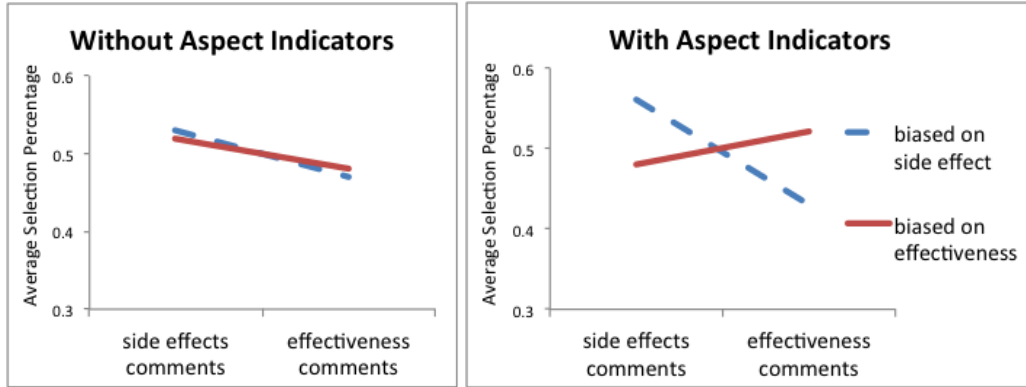
Figure 6.3: **Selection percentage of side effects/effectiveness comments for participants with different biased aspects**

biased aspect and comment aspect in the condition with aspect indicators, and found it to be significant ($\beta = 0.16, t(32) = 3.98, p < 0.01$). This result suggested that the manipulation successfully led to concerns (i.e., biased beliefs) focused on one of the two aspects, and the aspect indicator facilitated the selectivity.

In the remainder of the chapter, I will study the participants' information selection to understand: 1) Selective exposure tendency in the context of seeking medical opinions, and how it was mediated by disease risk and biased aspect (RQ1); and 2) How aspect indicators influenced selective exposure (RQ2). Lastly, I will examine whether selective exposure led to decision bias, and whether aspect indicators made a difference (RQ3).

6.3.1 Selective Exposure (RQ1 & RQ2)

For each participant, for each disease task he or she completed, I calculated the selection percentage of the four types of comment (attitude consistent/inconsistent, on side effects/effectiveness). I explored the information selectivity in detail by examining how the following independent variables mediated the selectivity of the four types of comments: disease risk (high/low), biased aspect (side effects/effectiveness), and the presence of aspect indicators (with/without). To this end, I ran a mixed-effect linear regression model on the selection percentages, by including comment attitude consis-

tency, comment aspect, disease risk, biased aspect, and presence of aspect indicators as fixed-effect independent variables. I found a significant three-way interaction among attitude consistency, biased aspect and disease risk ($\beta = 0.07, t(63) = 2.10, p = 0.04$), and a significant four-way interaction among attitude consistency, comment aspect, the presence of aspect indicators and disease risk ($\beta = 0.14, t(63) = 2.22, p = 0.03$) [1]. In the following sections, I interpret these interactive effects to uncover how these variables impacted the participants' selective exposure.

Effects of Biased Aspect and Disease Risk on Selective Exposure (RQ1)

The three-way interaction among attitude consistency, biased aspect and disease risk suggested that biased aspect affected selective exposure tendency differently for tasks regarding high-risk versus low-risk diseases. To illustrate the interaction, I plot the average selection percentages of each type of comment for decisions regarding low versus high-risk diseases separately in Figure 6.4.
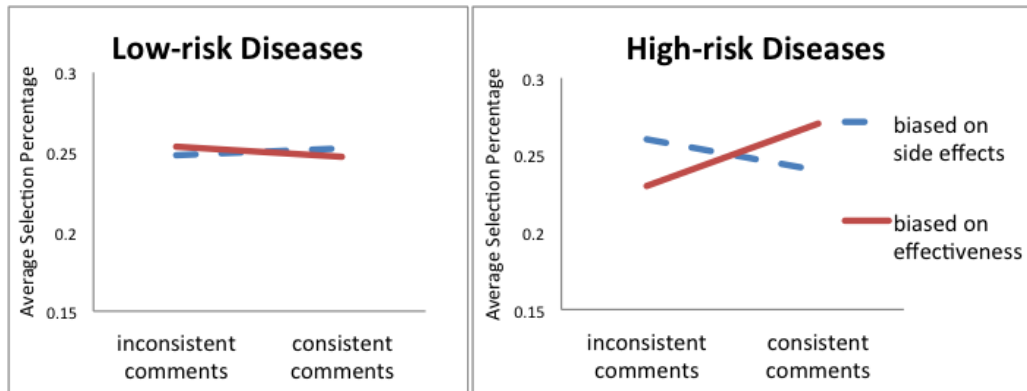
Figure 6.4: **Selection percentages of different comments by participants with biased beliefs on side effects or effectiveness**

The figures show that, when participants had biased concerns on the effectiveness of medication treating high-risk diseases, they exhibited pronounced selective exposure, by selecting more attitude consistent comments than inconsistent comments. They did not show such bias when making decisions

---

[1]The three-way interaction among biased aspect, comment aspect and the presence of aspect indicators was still significant.

104

for low-risk diseases, nor when they were only concerned about the side effects of medicines treating high-risk diseases. The conclusion is confirmed by the significant interaction between biased aspect and comment consistency for high-risk diseases ($\beta = 0.06, t(65) = 2.28, p = 0.02$), but not for low-risk diseases ($\beta = -0.01, t(65) = -0.50, p = 0.62$).

The result suggested that, when seeking medical opinions using this aggregator system, the participants' selective exposure was generally low, with one exception—when they were concerned about the effectiveness of treatment for high-risk diseases. I noted that this was likely the most threatening situation among all the scenarios. Not only that the information-seeking task regarding high-risk diseases was more likely to invoke threats than that of low-risk diseases, but also, according to previous work (Bergus et al., 2002), effectiveness is the more *critical aspect* than side effects in decision-making regarding high-risk disease. This is intuitive since it is vital to treat high-risk diseases, so patients are more likely to be tolerant of the side effects as long as the disease can be controlled. To verify that the two aspects may differ in criticality in medical decisions, I tested the information selectivity between comments concerning side-effects versus effectiveness. The intuition was that information-seekers would naturally look for more information regarding the more critical aspect for a multi-attribute decision.

I ran a mixed-effect linear regression on the selection percentage of side effect/ effectiveness comments (collapsing other variables, looking at only percentages of the two aspects) by including comment aspect, disease risk, and the presence of aspect indicators as fixed-effect independent variables. I found a significant three-way interaction among them ($\beta = 0.21, t(64) = 3.93, p < 0.01$). In Figure 6.5, I illustrate this three-way interaction by plotting the average selection percentages of each type of comments.

It suggested that, with the presence of aspect indicators, participants sought more effectiveness related information for high-risk diseases, but more side effects related comments for low-risk diseases. The conclusion was confirmed by the significant two-way interaction between comment aspect and disease risk found in the conditions with aspect indicators ($\beta = 0.10, t(32) = 2.35, p = 0.02$). The observation was consistent with what was found in Bergus et al. [169], suggesting that effectiveness is the more critical aspect than side effects when selecting treatment for high-risk diseases, while this

105
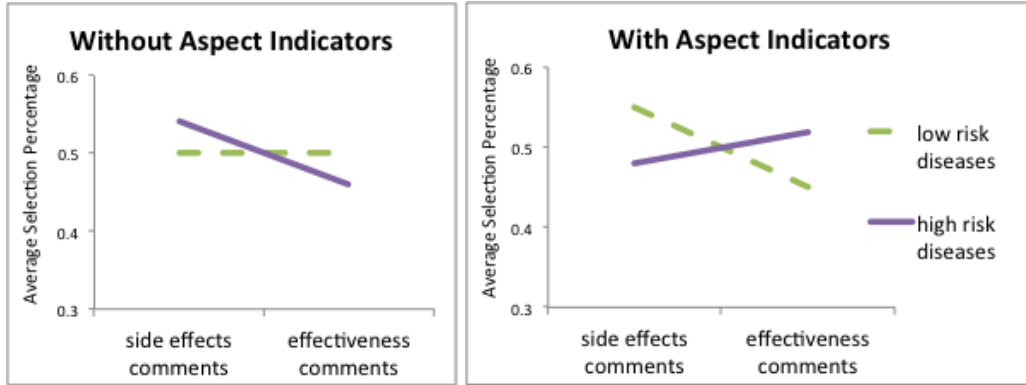
Figure 6.5: **Selection percentages of side effects/effectiveness related comments for treating high versus low risk diseases**

difference in criticality tends to be moderated, even reversed for low-risk diseases.

Taking together, the above results presented a warning that, when seeking medical opinions, selective exposure may become more pronounced in more critical situations, such as when concerning about the effectiveness of treating high-risk diseases. This echoed results in a study conducted by Jonas et al. [167], finding that in more threatening situations, specifically when mortality is a salient possibility, medical decision-makers tend to show pronounced selective exposure. These results, again, showed that heightened threat increases the selective exposure tendency, and is especially problematic in health-related contexts, as people appear to be most vulnerable to biased information seeking in situations where decision errors are especially disastrous.

Note that I did not find the presence of aspect indicators to mediate the above three-way interaction (four-way interaction $\beta = -0.02, t(63) = -0.35, p = 0.73$), suggesting that the overall selective exposure tendency was not changed by the presence of aspect indicators. However, the four-way interaction found significance among attitude consistency, comment aspect, the presence of aspect indicators and disease risk suggesting that the aspect indicator might have impacted selective exposure to information regarding the two aspects in different ways. I unpack the four-way interaction in the next section.

106

Effects of Aspect Indicators on Selective Exposure (RQ2)

To begin with, I tested the three-way interaction among attitude consistency, comment aspect and presence of aspect indicators for tasks regarding high-risk and low-risk disease separately. I found it to be significant ($\beta = 0.12, t(64) = 2.51, p = 0.01$) for high-risk disease tasks, but not for low-risk disease tasks ($\beta = -0.02, t(64) = -0.53, p = 0.59$). It suggested that the aspect indicators might have changed selective exposure for tasks regarding high-risk diseases, but not for information in either aspect of low-risk diseases, as further illustrated by the descriptive results shown in Figure 6.6 (i.e., slops did not change). Therefore in the rest of the section, I focus on the high-risk disease tasks.
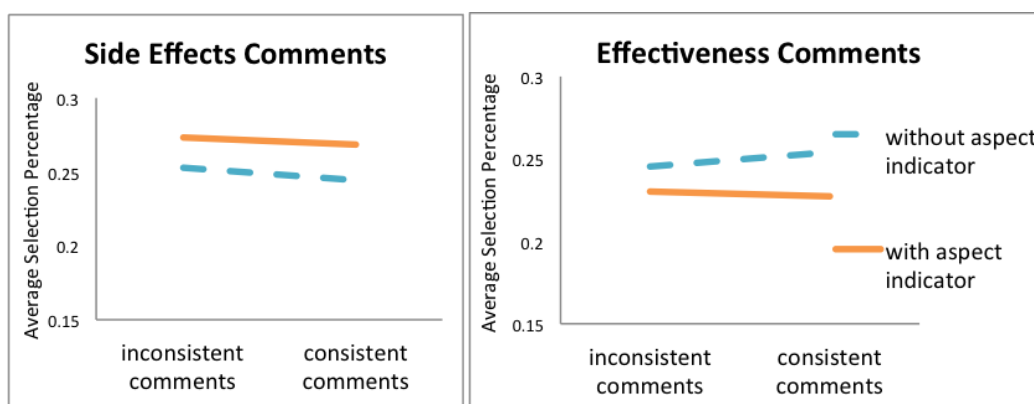
Figure 6.6: **Average selection percentage of different types of comment in low-risk disease tasks**

In Figure 6.7, I plot the average selection percentages of different types of comments for high-risk disease tasks. It shows that, when seeking information for treating high-risk diseases, the presence of aspect indicators increased the selective exposure tendency for seeking effectiveness related information, but decreased it for seeking side effects related information. In the following two subsections, I discuss the implications for the two conditions—biased on effectiveness or side effects, respectively.

Biased on Effectiveness Aspect in High-Risk Disease Tasks

The decrease of selective exposure to side effects related information could be good news for the most concerning case—when users were concerned about the effectiveness of treating high-risk diseases, which I have shown to incur
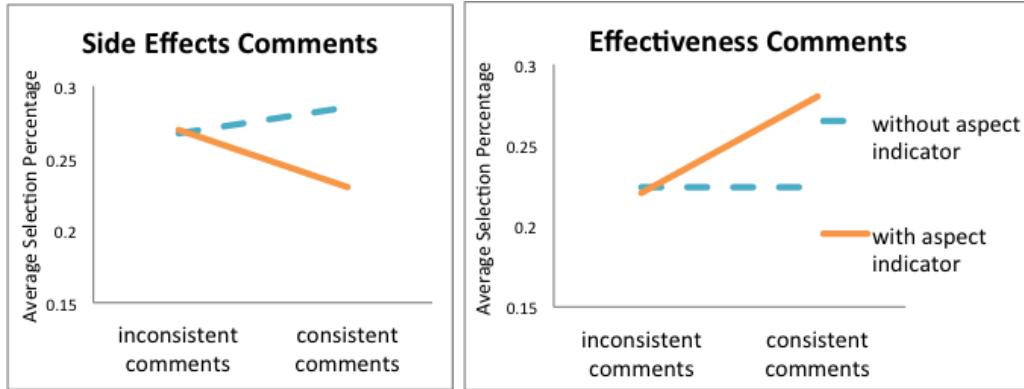
Figure 6.7: **Average selection percentage of different types of comment in high-risk disease tasks**

the most pronounced selective exposure. With aspect indicators, the users' selective exposure in the "unbiased aspect", i.e., side effects, could have been reduced. To verify, I tested the interactive effect between attitude consistency, comment aspect and aspect indicators in this situation, and found it to be significant ($\beta = 0.18, t(64) = 2.77, p < 0.01$). I plot the average selection percentages of different types of comments in the particular situation in Figure 6.8.

The figure demonstrates that, in the control condition where participants could not discriminate between comments with different focused aspects, they exhibited selective exposure for seeking both effectiveness and side effects related comments, even though they only had biased beliefs concerning the effectiveness. In the experimental condition, the aspect indicators helped participants to distinguish between the two kinds of comments, and hence "nudged" them to only show selective exposure on the effectiveness aspect, but not on the side effect aspect. To verify that selective exposure in seeking side effects related comments was decreased, I tested the interactive effect of attitude consistency and the presence of aspect indicators on the selection percentage of side-effects related comments in the situation, and found it to be marginally significant ($\beta = -0.07, t(32) = -2.02, p = 0.08$).

This is an encouraging finding. As I hypothesized, distinguishing the focused issue aspects could encourage people to make more deliberative choices of information, and alleviate selective exposure in aspects where they do not have biased beliefs. This is especially important for multi-attribute decision-

Figure 6.8: **Average selection percentages of different comments in the situation of "biased on effectiveness for treating high-risk diseases"**

making where accurate assessment of each aspects is required [170]. For example, imagine someone with negative opinions on the effectiveness of a treatment: if s/he indiscriminately seeks biased information for all aspects, s/he may end up concluding that the treatment is both ineffective and can cause bad side effects. Aspect indicators may potentially mitigate the problem, which I will further examine in the decision outcome section.

Biased on Side-Effects Aspect in High-Risk Disease Tasks

I also conducted the same test on the interactive effect between attitude consistency, comment aspect and aspect indicators in the situation where participants had biased beliefs in the side effects of high-risk diseases, and found no significant result ($\beta = 0.05, t(64) = 0.76, p = 0.45$). It indicated that in this situation, aspect indicators did not significantly change the selective exposure tendency. On the one hand, this is good news that it did not increase the participants' selective exposure in either aspect. On the other hand, it did not fulfill the expectation that aspect indicators would reduce selective exposure in the unbiased aspect, i.e., effectiveness in this situation. To verify, I tested the interactive effect between the presence of aspect indicators and attitude consistency on selection of effectiveness related comments, and found no significant result ($\beta = 0.02, t(31) = 0.42, p = 0.68$). One possible explanation is that since effectiveness for treating high-risk diseases tends to be the critical aspect, aspect indicators would be less effective in reducing selective exposure even though it was the unbiased aspect. In other

109

words, whether aspect indicators could work to reduce selective exposure in the unbiased aspect may be further mediated by the aspect criticality in the decision context.

In short, in this section I found evidence that *aspect indicators could reduce selective exposure in aspect one did not have biased beliefs, at least when the unbiased aspect was not the critical aspect of the decision.* I also found that, in general, *aspect indicators had a stronger impact on information selectivity for tasks regarding high-risk than low-risk medical information seeking contexts.* This is consistent with the prediction that, since disease risk is directly relevant to outcome importance, a contributor to accuracy motivation, it may mediate the effectiveness of aspect indicators as it is designed to be a central nudge.

6.3.2 Decision Outcome (RQ3)

In this section, I explore the outcome of selective exposure on decision bias, and whether aspect indicators improved the decision. As discussed earlier, I measured decision outcome in the post-task survey by two 5-point Likert-scale items, and created the decision index by averaging the ratings of the two. The value reflected the preference for the alternative over the current option, i.e., decision bias consistent with the pre-existing attitude, with higher value indicating higher preference. For quality control reason, I excluded 10% of the results where the ratings given to the two questions differed by larger than 2.

To verify that the participants' decision outcomes were influenced by their selective exposure, I started by testing the effect of selective exposure on decision outcomes. For each task performed by each participant, I quantified the extent of selective exposure, by the number of attitude consistent comments minus that of attitude inconsistent comments selected, and tested its main effect on the decision index. As expected, I found a significant positive effect of selective exposure on the decision outcomes ($\beta = 1.06, t(66) = 4.62, p < 0.01$), i.e., the preference for the alternative option. It confirmed that the participants' decision bias was closely associated with the extent of selective exposure in their information seeking process. I ran a mixed

effect linear regression model on the decision index, by including disease risk, biased aspect and presence of aspect indicators as fixed-effect independent variables. I found a significant three-way interaction among all of them ($\beta = -1.32, t(64) = -2.06, p = 0.04$), and a two-way interaction between disease risk and biased aspect ($\beta = 0.91, t(64) = 2.00, p = 0.05$).

I plot the decision bias in different situations in Figure 6.9. Consistent with the results of selective exposure, it shows that in the control condition, the situation with the most pronounced selective exposure—having biased beliefs in effectiveness of treating high-risk disease—led to the strongest decision bias. However, the presence of aspect indicators mitigated the problem. To further verify, I tested the interactive effect between disease risk and biased aspect separately for the control and experimental conditions. As expected, I found it to be significant without aspect indicators (=-0.94, t(62)=2.12,p=0.03), but not significant when aspect indicators were presented (=-0.40, t(65)=-0.86, p=0.40).

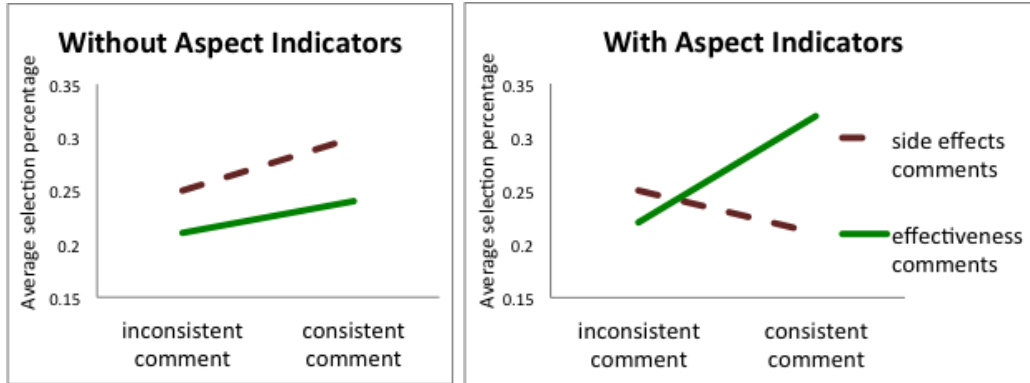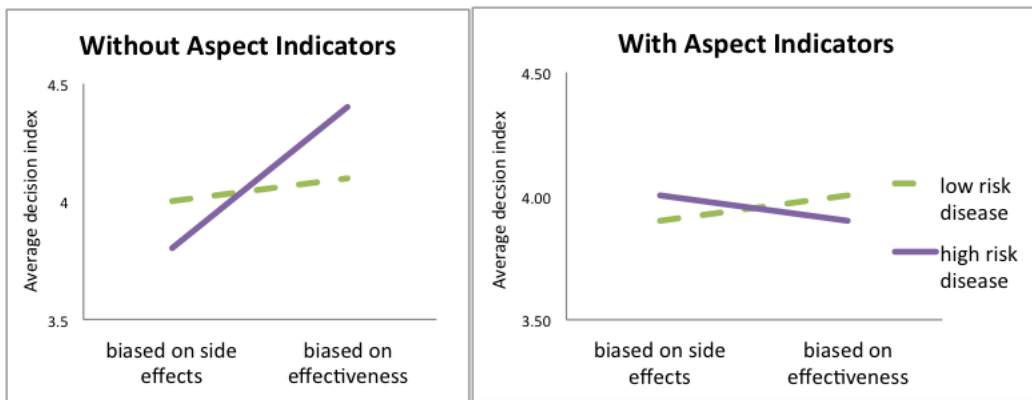Figure 6.9: **Average selection percentages of different comments in the situation of "biased on effectiveness for treating high-risk diseases"**

To summarize, I found evidence that *the pronounced selective exposure in threatening situation in medical opinion seeking could lead to decision bias.* However, *aspect indicators were effective in mitigating the decision bias by reducing selective exposure in the "unbiased aspect".*

6.4   Discussion

In this experiment, I used an interface similar to the ones used in Study 2 and Study 3 (Chapter 4 and Chapter 5), but applied it to a different task domain—consumers seeking social opinions to make medical decisions. By comparing selective exposure in different contexts, specifically, concerned with varied aspects for disease with high or low level of risk, I showed that selective exposure became more pronounced in the more critical and threatening situation, which also led to higher decision bias. I also showed that aspect indicators, by distinguishing information that focuses on different issue aspects, was able to mitigate the problem, by reducing selective exposure to information on "unbiased aspect" where users did not have biased beliefs.

A potential implication of the study is that health-related information seeking may create situations where accuracy motivation and threat (i.e., defense motivation related factor) both get heightened. On the one hand, there would be high outcome importance related to one's health status, even life, to accentuate accuracy motivation. On the other hand, the situation would alarmingly provoke threat and anxiety, as the possibilities of risk and even mortality become salient [167]. A possible outcome, as the results suggest, is that in critical medical decisions heightened threat can play a more dominant role and thus increase the selective exposure tendency. Meanwhile, due to the high accuracy motivation, users in such situations also tend to react to "central nudges" that can help make more deliberative choices of information. In contrast, users who are facing low-risk decisions may have relatively lower accuracy motivation, and thus are less vigilant in information seeking to make conscious choices of information based on its content.

However, aspect indicators should still be used with caution, as their effectiveness may depend on the criticality of the aspect. In this study, their effectiveness on moderating selective exposure in unbiased aspect seemed to be weaker when the unbiased aspect happens to be the critical aspect of the decision. I should point out, however, that more research is needed to validate this conclusion. It is possible that the weakened effect was due to the tightened anxiety associated with seeking effective treatment for high-risk diseases, and we do not know whether it applies to other contexts, e.g., in political debates. Moreover, the selective exposure tendency in this situa-

tion, where one was only concerned about side effects of a life-saving drug, seemed to be low in general, and that could also have led to the absence of additional positive effects of aspect indicators.

As with previous experiments, while I chose to study an "indicator" design, I expect the result to inform about the benefits of a category of interface features that categorize information by aspects. Two of the most common examples are faceted interface and tags (i.e., the keywords of the message, and thus its focus). This study suggests that these design features can potentially help improve deliberation and unbiased learning, and recommends incorporating them in deliberation platform and information systems that present diverse viewpoints.

The fundamental idea of aspect indicators as a central nudge is to provide more fine-grained division of the issue and distinguish aspects where users may face less cognitive dissonance, and thus potentially more open to challenging views. A more advanced version of the idea would be to highlight the "unbiased aspects". For example, in a political forum, a feature that enables users to see "topics you both agree" can potentially reduce the barriers of interactions between users with different opinions or ideology. Another similar idea is to reconstruct messages that support a disagreeable position by aspects, and bring forward the non-contradicting aspects before introducing the more attitude-challenging ones, similar to the de-biasing techniques recommended in [130].

To implement these design ideas, existing NLP techniques that cluster content or extract aspect would be useful (e.g., [27, 100]), but may not be sufficient. This study suggests that, to develop more accurate nudges, beyond "aspects in the information", we should also explore ways to understand "aspects in the mind", to accurately infer on which aspect a user may or may not have biased beliefs, and which aspect may be considered critical.

Last but not the least, this study contributes to the literature on biases in consumer online health information seeking. Consistent with previous research [31, 50, 109], I found that health information seekers tend to seek more confirmatory information in anxiety prone situation such as when one is concerned about a life-threatening condition. Given that in this study I used a hypothetical scenario, in real patients' information seeking contexts the

perceived threat can be further heightened. So selective exposure could be even more pronounced than what was observed in the study. Therefore, debiasing techniques should be taken into consideration when designing medical decision-support systems. This study also suggests that, while medical decision is typically a multi-attribute decision, information seekers may not be able to recognize or explicitly seek information to accurately assess all the different attributes. Aspect indicators, or other kinds of design feature that encourage users to explore each attribute separately and thoroughly, may facilitate more informed decision-making. To support assembling these assessments, DSS should also consider that weights given to different aspects may vary depending on the decision contexts and user profiles.

Chapter 7

EXPLORING MODERATORS OF SELECTIVE EXPOSURE AND OPINION GROUP BEHAVIOR DIFFERENCE

In the last decade, there has been a growing interest in empirically examining the segregation and polarization of ideological and opinion groups on social media. The concern, needless to say, is rooted in selective exposure—driven to seek confirmation from likeminded others, people may choose to disproportionally connect and interact with others who share similar views. In previous chapters, I show that moderators lead to differences in selective exposure at the individual level. In this chapter, I explore whether group attributes associated with moderators could lead to group differences in selective exposure tendency—the in-group bias. As discussed in the related work section, several studies have shown that some opinion or ideology groups tend to show stronger in-group bias on social media by more actively connecting and interacting with in-group members than others, segregating them in isolated echo chambers [66, 71, 117, 172].

Defending the role of selective exposure, some communication researchers associate selective exposure with increasing political interest, participation, and expressiveness, due to the positive social proof and strengthened confidence [11, 173]. Consistent with this view, empirical studies on group segregation on social media often found the more densely connected group to be more active [52, 53, 174]. It is therefore possible that the group difference in selective exposure may drive group difference in expressiveness—members of one group become more active and expressive about their opinions, by engaging in active discussions with like-minded others, where their positions are constantly reinforced.

Meanwhile, group difference in expressiveness has been discussed as a critical warning for biased presence of different groups. Not only does it impact the perception and exposure for users of the platform, but also, it may introduce biases to analytics that use social media data to understand and

monitor social opinions. For example, Mustafaraj et al. [174] pointed to the existence of a "vocal minority" and a "silent majority" in Twitter discussions, so that the vocal group might appear to be over-represented in the datasets. They also further showed that the two groups differ significantly in their tweeting behavior, hence the content they created, with the vocal group engaging in more activities that aim to broaden the group impact, such as mentioning others and using hashtags. With the results, they warned to exercise cautions against aggregating data to build predictive models regardless of group differences in expressiveness. Some scholars also warned that because expressiveness tends to correlate with categorical attributes, i.e., the more involved or more impacted groups tend to be more active (e.g., consider the more prevailing negative complaints on some review websites), it would create inherent sampling bias for using social media data to study human behaviors [175].

While many factors can drive group differences in expressiveness, in this work, I explore the association between group difference in selective exposure tendency, which has also been called in-group bias by researchers looking at the problem at the group level, and group difference in activeness and expressiveness. Note that I do not claim the causal relationship to be strictly one-way, as social selection and social influence often happen in both ways [176].

Specifically, I examine the case when there are a numerical majority group and a numerical minority opinion group. On the one hand, social science research on in-group process robustly shows that people of the minority group tend to exhibit stronger in-group bias (see review in [122]). This is often explained by the threat and insecurity that minorities face, driving them to actively look to connect with people of the same category, seeking and providing support to each other [123]. Some also argue that by actively engaging in selective exposure, and as a result, shielding oneself in echo chambers of reinforcing information, is key to the survival and development of minority social groups [177].

On the other hand, theorists of "spiral of silence" predict the contrary [178]. The theory suggests that opinion minorities may refrain from expressing their position due to the fear of social isolation. As a result, the group may fall silent in spite of the actual distribution in the society. However, empirical

116

studies that attempted to prove the spiral of silence have encountered many failures (see review by Scheufle and Moy [179]). Researchers then pointed to the moderating effect of attitude certainty, i.e., the existence of a "hardcore minority" who would speak out regardless of the social environments.

In this chapter, I will empirically examine whether the minority group would fall silent or become louder in expressing their opinions over time. Although theories provide mixed suggestions, I note that Twitter is a social-network based system where people can actively curate their network and their information feeds. So compared to common platform type of systems such as discussion forums, Twitter may give more leeway to selective exposure to happen, and therefore, creating echo chamber where one would fear less about social isolation thus experience less need for silence.

Another important goal of this study is to examine the asymmetries of "peripheral nudges"—group differences in the collective use of social features. As demonstrated in Study 3 (Chapter 5), social peripheral features (e.g., popularity, endorsement and source competence cues) for information or information sources can nudge users to attend to the information, and also promote more favorable perception of the information, regardless of its position. However, the problem is that, in most real-world systems, peripheral nudges are not created equally for different viewpoints. Given that social features such as "like," sharing and endorsement, are mostly to express support, we can view the peripheral nudges as outcomes of the collective use of the social feature by people of the same opinion group. Therefore, if certain opinion groups tend to use a social feature more actively than others, that would create asymmetrical nudges for that particular position.

Of course, the inequality in peripheral nudge is only one of the many problems with asymmetric use of social features between different opinion groups. Given that many algorithms make use of these social features, it could also asymmetrically impact the algorithm output of different viewpoints. Moreover, social sharing features also asymmetrically increase the numbers of appearance of different viewpoints. This not only influences the reachability and impact of different viewpoints on the platform, but also introduces biases to the presence of different opinions in the data, creating a potential problem for social media analytics.

In this work, I study the group differences in the collective use of one social feature on Twitter—the retweeting button. Specifically, I study the differences between a numerical majority and a numerical minority group. I focus on the asymmetric numbers of retweeting that the two groups received, as the source of biases that can potentially *amplify* the presence of certain groups over others, including peripheral nudge for their opinions, number of presence, and algorithm ranking, etc. Previous research has highlighted that the collective use of social features may create undesired inequality in information cascade [146, 180], leading to greater inequality between popular and unpopular items. Given that the majority group has a larger user base, thus their views are considered more popular on the platform, we may expect that the collective use of social features would give the majority group further advantages. In this work, I study the bias, size and changes over time of this asymmetrical outcome of the collective use of social features.

7.1 Overview of the Case Study

To explore these issues, I present a case study of the 10-month Twitter discussion on Edward Snowden, the former NSA subcontractor who made global headlines by leaking secret documents that expose the NSA's global surveillance program. Especially in the United States, Edward Snowden is a controversial figure who fueled much debate on government surveillance and information privacy. He has been considered a "hero" and a "patriot" by many, but has also called a "traitor" by some for the potential threat he posed to the national security. Although media polls in the US showed mixed results of public opinions on Snowden [181, 182, 183], there has been anecdotal evidence that Twitter users dominantly lean towards pro-Snowden [184]. Strikingly, when I looked at the top 100 most retweeted tweets in the dataset, only one of them expressed minor doubts about Snowden, but almost 40 of them expressed strong support for him. I also expect the Snowden discussions to be suitable for group behavior changes given the continuous interests it generated over time.

By investigating the dataset, my goal is to explore how the numerical majority (pro-Snowden) and numerical minority (anti-Snowden) groups differ

Table 7.1: **Research questions and findings in Study 5 (Chapter 7)**

	Research question	**Findings**
RQ1	How did the majority and minority groups' activity level change?	Comparatively, the minority group became more active over time.
RQ2	How did the two groups differ in expressiveness (content characteristics)? How did they change?	The minority group tended to produce more opinionated content than the majority group, and over time became increasingly so.
RQ3	How are the group differences in activeness and expressiveness associated with varied in-group bias?	Evidence was found that the minority group showed stronger in-group bias, and over time focused even more on interpersonal interactions, including conversations with like-minded others, which was a main reason that they tweeted more actively and expressively over time.
RQ4	How did the majority and minority group differ in collective use of the retweet button? How did it change?	The majority group received more retweets, which amplified the presence of and peripheral nudge for their positions. The amplification effect increased over time, which was found to be associated with the active and persistent role that the pro-Snowden opinion leaders play.

in their activeness and expressiveness, and how the difference, if any, is associated with their difference in in-group bias. I also study whether their collective use of social features creates asymmetrical results that amplify the presence of and peripheral nudge towards one of the opinion groups. Although previous research compared the in-group bias of opposing opinion groups on social media [66, 71, 117, 172], the case of numerical majority and minority groups has not yet been explored. Specifically, I ask the following research questions in Table 7.1.

7.2   Dataset

By using Twitter Streaming API, we collected all publicly available tweets containing "#snowden" (case insensitive). The collection started on July 6, 2013, one month after Snowden's first public appearance, and ended on

April 25, covering a period of 42 weeks. Including both original tweets and retweets, the dataset included 1.06 million English tweets. We used the Python NLTK package to exclude non-English tweets as we chose to focus on the English-speaking opinion space.

The dataset contained meta-data indicating whether a tweet is a retweet and if so, the ID of the original tweet. The meta-data is automatically generated when user used the "retweet" button. There are other ways people retweet, e.g., by manually copying and adding "RT," "retweet," and so forth. However, we counted tweets that contain these markers and only 7.9% of them are not shared through the retweet button. This number is significantly lower than what was reported in earlier studies (e.g., [185]), which were conducted not long after Twitter introduced the retweet button. This result shows users' adaption to platform-provided features. Given that we are interested in the effect of the retweet button, we focus on automatically generated retweets only, and treat hand-copied ones as original tweets—to some extent, they are similar to tweets posted to share information. Excluding retweets, the dataset includes 440k original tweets. We also used the meta-data information to track how many times an original tweet was retweeted.

Figure 7.1 shows the volume of tweets by week. The trend is generally consistent with Google Trends data showing the number of times "Snowden" was searched. A burst of attention following his first media appearance, which decreased rapidly in the first couple of months. From October 2013 onwards tweeting activities were generally stable. However, there were "local peaks" following significant events such as Snowden's Christmas speech and the SXSW speech, etc., showing that there were continuous topical interests.

7.2.1 User Classification

Inferring Twitter users' opinions is well-known to be a challenging task [80, 172]. After experimenting with different approaches, I chose to follow [71] by using a variation of a Raghavan's label propagation algorithm with manually labeled "seeds." The method is based on the robust observation that retweeting is an endorsing activity [186] and one's position can be inferred
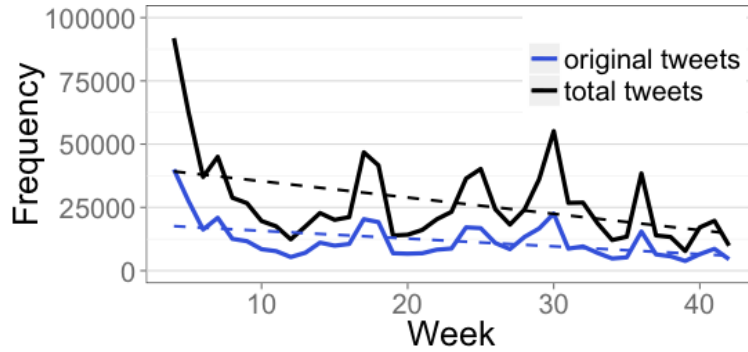
Figure 7.1: **Total and original tweets volume by week**

based on whom or what one retweets [187, 188, 117, 189].

For the retweeting network, I treat each user as a node, and each time a user retweeting another user as an out-going edge. The algorithm works in iterations: In each iteration, I start with a pro-Snowden cluster and an anti-Snowden cluster of nodes. Then I iterate every node in the dataset to assign or re-assign each node to the cluster to which it has more out-going edges. When they are equal (but not 0), I assign it to the minority group to compensate for the smaller group size. I use the result of the current iteration to seed the next iteration, and I use a hand-labeled sample with confidently known positions to seed the initial iteration. I always fix the cluster of the initial seeds, and give outgoing edges to them a slightly higher weight by setting weight for edges to non-seeds at 0.8.

To obtain a sample of seeds, I crowdsourced opinion labels for 3% of random sample of daily original tweets. I chose to label tweets instead of users because it was easier and more suitable for crowdsourcing tasks. Random sampling also made it more likely to include active users, which are better positioned for seeding the network clustering algorithm. I had three Turkers to label each tweet. They were asked to label a tweet only if they could confidently judge its position or otherwise leave it in the "unknown" category. If the three selected different labels thus did not reach a majority vote, we recruited 2 more Turkers to label the same tweet. I worked to verify labels for tweets that did not get high agreement. Eventually, 23.5% of sampled tweets were confidently judged with positions. With this "conservative" labeling schema targeting high precision, I classified authors of tweets labeled as pro or anti to be pro-Snowden or anti-Snowden seeds. Only 0.6% of them

appeared in both groups, and I manually examined their tweets to decide their positions, if possible. I ended up with a sample of 519 anti-Snowden and 1,922 pro-Snowden seeds.

I used the labeled sample to seed the above-mentioned algorithm, which reached stability at the 56th iteration. I focused on users who actively participated in the discussion, i.e., those who produced at least one original tweet. I was able to identify 27,119 pro-Snowden and 2,400 anti-Snowden users. Their tweets accounted for 69.1% of the whole original tweets dataset. The cumulative distributions of the number of tweets of the two groups are fairly similar, with a few more extremely active users in the pro-Snowden group.

For evaluation, I randomly sampled 200 pro-Snowden users and 100 anti-Snowden users from our identified users. I pooled them together and retrieved all their tweets. Two other researchers rated whether each person leans towards pro or anti-Snowden, or they could not judge because either the person had too few tweets or was only sharing factual information. Table 7.2 presents the evaluation results. In general, the classification method gave satisfactory results, with around 90% accuracy for each category.

Group	agree	not agree	cannot judge
pro	90.7%	1.3%	8.0%
anti	89.0%	4.0%	7.0%

Table 7.2: **User classification evaluation results**

7.2.2   Top Influencers

In some analyses, I explore the role that opinion leaders play—people who exert highest influence within each group. Cha et al. [190] concluded that the number of retweets a user received can reflect a user's topical influence. I therefore chose to identify the *top influencers* based on the total number of retweets they received. For each opinion group, I identify users above the 50% of the cumulative distribution of number of retweets received to be the top influencers—this was to account for the fact that the retweet distribution of pro-Snowden group has a heavier tail. 14 anti-Snowden users and 44 pro-Snowden users were identified to be top influencers. Note that

this method may leave room for improvements. However, having tried more sophisticated methods (e.g., also considering mentioning), the results are generally correlated and do not change the conclusions. Hence we present the results by applying the simplest method.

7.3 Results

I will first study the group differences and changes in activeness (RQ1) and expressiveness (RQ2). I will then look at the group difference in selective exposure tendency, i.e., in-group biases, in inter-personal interactions, and explore its association with the group difference in activeness and expressiveness (RQ3). In the last part of this section, I will examine the group difference in collective use of the retweeting feature, and how it changes over time (RQ4).

7.3.1 Change of Activity Levels (RQ1)

I start by seeking answers to RQ1—the changes of the activity levels of the anti-Snowden group. I am interested in knowing whether the minority group's presence increased or diminished in the opinion space. I choose to look at the relative size of *tweets volume* and *participating users* (who posted original messages) from the anti-Snowden group. The overall activity change has been discussed in the dataset overview section (Figure 7.1). In Figure 7.2, I present the relative proportion of tweets from the anti-Snowden group among all the identified users. Results are calculated by the unit of week, i.e., how many users posted tweets in each week. Varying from week to week, there were 5% to 45% of users among all the identified users participating.

Following [191], I study temporal changes by the linear fitted *trend line* with weekly data points, which allows me to compare the temporal changes of groups by testing the difference between their slopes (with t-test for the coefficient difference divided by the pooled standard error). Both the proportion of tweets volume ($t(40) = 3.41, p < 0.01$) and the size of users from

the anti group ($t(40) = 2.62, p = 0.01$) show slopes significantly higher than 0, and the difference between the two is significant ($t(80) = 1.94, p = 0.05$). This suggests that over time the anti-Snowden group's tweeting activities were better sustained than the pro-Snowden group. Also, on average, the anti-Snowden users became more active than those in the pro group in producing original tweets, suggested by the higher proportion of tweets than the proportion of users. This is an interesting observation suggesting that, instead of silencing themselves, the minority group became comparatively more active and increased their presence in the opinion space (of original messages).
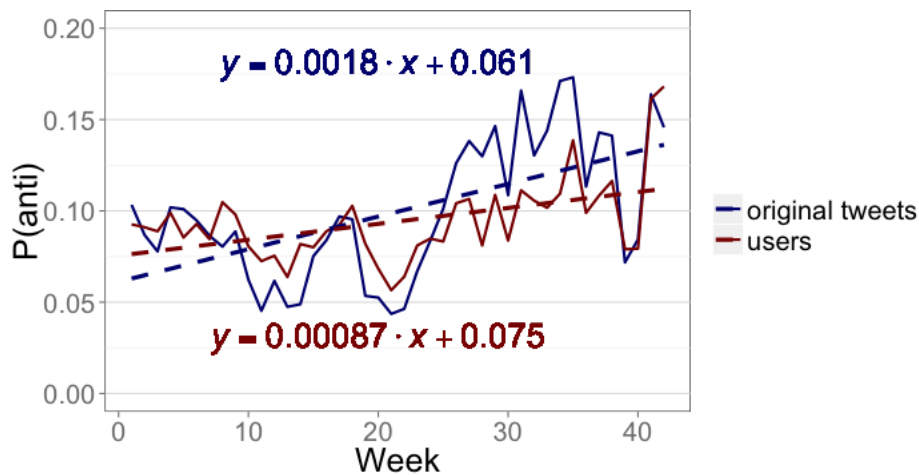
Figure 7.2: **Percentages of tweets and users from the anti-Snowden group by week**

For all the temporal changes I study in this chapter, I need to exclude an alternative explanation that the changes could be caused by a significant change in the population of the two groups, instead of group behavioral differences. To do so, I look at the distribution of the weeks members joined the discussions (the first time one tweeted #Snowden). A K-S test comparing the distributions of the two groups showed no significant difference (p=0.58). Also, I looked at the "top active users" above 50% of the cumulative distribution of the total number of tweets for each group, 92.3% in the anti and 95.6% in the pro-group were already tweeting in the first half period, and 93.6% in anti and 93.9% in the pro-group were still doing so in the latter half, suggesting that the highly active users were generally persistent. Therefore, I can reasonably conclude that the temporal changes I observe should not be

due to significant differences in the population changes between the groups.

7.3.2 Content Characteristics (Expressiveness) (RQ2)

To answer RQ2, I am interested in whether the two groups exhibited differences in their expressiveness, and thus produced contents with different characteristics. Identifying such differences would have important implications for social media analytics using content-based measurement to understand social opinions. Twitter provides widely used "markers" to inquire about communication patterns at the aggregate level. For events related discussions, Bruns et al. [192, 193] suggest that using two metrics—the percentage of original tweets containing URL, and the percentage of retweets among all tweets—one can classify if the discussions focus more on information sharing (high URL and retweets percentage), or more on making original commentary (low URL and retweets percentage)—the later is direct relevant to the tendency of expressiveness I want to study. By sampling Twitter discussions from 40 hashtags related to political, social and entertainment events, they demonstrated that they distinctively fell into one of the two clusters [192].
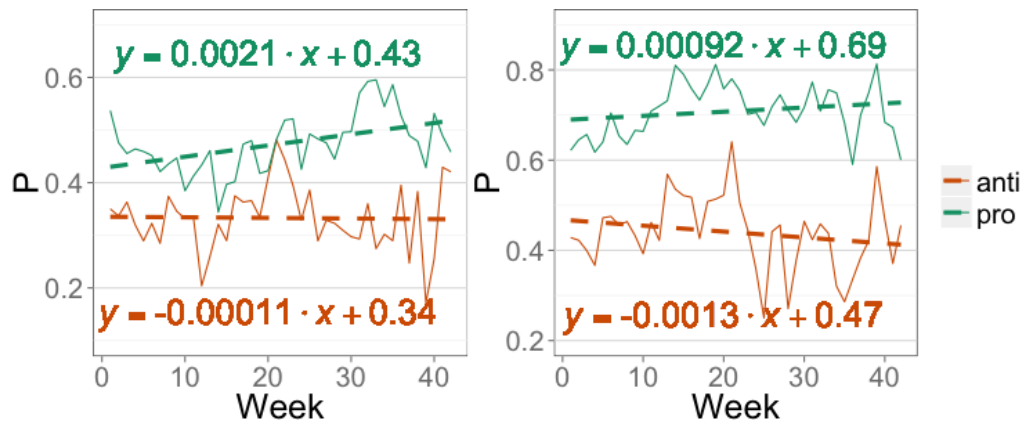
Figure 7.3: **(left) Percentage of retweets among all tweets by week; (right) Percentage of original tweets containing URL by week**

Figure 7.3 shows the percentage of retweets (left) and tweets with URL (right) by week for each group. There are two main observations. First, the anti-Snowden group focused more on making original commentary, characterized by lower percentage of URL shared and lower percentage of retweets.

This may resemble an *"audiencing"* experience where people watch and comment on unfolding events. [192] showed that tweets related to media events, such as political elections, where Twitter acts as a discussion back channel, often collectively exhibit such patterns. The pro-Snowden group appears to have centered more on sharing information, with higher percentage of URLs shared and a higher percentage of retweets. It may represent a *"gate-watching"* practice, with desire to disseminate key information, but limited interest in posting original comments. Second, we found that the differences became even more distinct over time, with the anti group becoming more commentary oriented, while the pro group focusing more on disseminating information. Comparison between the coefficients of the trend lines of the two groups is significant for retweets percentage ($t(80) = 2.19, p < 0.03$), and marginally significant for URL percentage ($t(80) = 1.80, p < 0.07$).

The above results imply that the anti-Snowden group became more active and also possibly more expressive over time. Imagine using a "pooled data" approach to study opinion changes on Snowden, one may draw the conclusion that the climate has turned against Snowden—however, could the behavioral changes be driven by other reasons? I highlight one important factor to consider—differences in social interactions, including in-group and cross-group interactions. Observing that 52.4% tweets from anti-Snowden group and 40.8% tweets from pro-Snowden group contained at least one mentioning, which signals inter-personal interactions, I note that social interaction itself is a main source of content in the Snowden Twitter discussion. Moreover, I argue that group differences in social interactions may construct dissimilar social environments, such as echo chambers, that could systematically drive group behavior changes. To shed light on such underlying mechanism, I study the differences in in-group and out-group interactions between anti and pro-Snowden groups in the next section.

7.3.3 In-group and Out-group Interactions (RQ3)

I will answer RQ3 by studying in-group (mentioning opinion-similar users) and out-group (mentioning opinion-different users) interactions. While mentioning can be used for different purposes, e.g., to reply, to initiate conversa-

126

tions, or to refer to another user, I do not differentiate but generally consider them as signaling interactions at the interpersonal level.

In Table 7.3 I present the number of total in- and out-group mentioning degree. For tweets that contain multiple mentions, we count them as multiple instances. I compare the ratio between the observed mentioning degree to the *expected degree* in the mentioning network, calculated as follow:

$$D[i->j] = d_i \cdot \frac{U_j}{U_{anti} + U_{pro}} \tag{7.1}$$

where $i, j = anti$ or pro. d_i is the total number of mentioning from group i to all the users we included in the analysis, and U_j is the total number of users in group j.

	N		Ratio to Expected Value	
	$\rightarrow$ Anti	$\rightarrow$ Pro	$\rightarrow$ Anti	$\rightarrow$ Pro
Anti	4639	7797	**4.59**	0.68
Pro	5150	71691	0.82	**1.02**

Table 7.3: **Total mentioning degree between and within groups. Ratio to expected values for in-group interactions are in bold.**

As shown in Table 7.3, I found that the minority group had more inter-group interactions than the majority group. However, considering the large difference in the group size, the observed within-group interactions in anti-Snowden group is still significantly higher than the expected value, and the ratio is much higher than that of the pro-Snowden group. This suggests that both groups exhibited ingroup bias in their mentioning behaviors, with much higher bias in the minority opinion group.

I then examined the temporal changes of in-group and out-group interactions. I first looked at the average total mentioning degree by week (averaging by people who produced original tweets). As shown in Figure 7.4 (left), over time mentioning activities increased in the anti-Snowden group, but decreased in the pro-Snowden group (the difference is significant at $t(80) = 2.10, p = 0.04$). I then examined, for those mentioning others, their average percentage of in-group mentioning and out-group mentioning. As shown in Figure 7.4 (middle), over time the in-group mentioning increased for anti-Snowden users, but decreased for pro-Snowden users, with the two
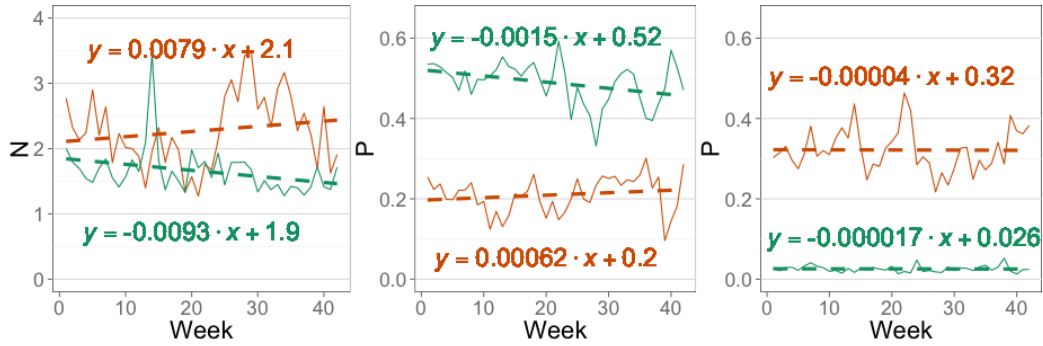
Figure 7.4: **(left) Average mentioning degree; (middle) Average in-group mentioning percentage by person; (right) Average out-group mentioning by person for pro (green) and anti-Snowden (red) groups**

trend lines significantly differed ($t(40) = 2.52, p = 0.01$). In contrast, as shown in Figure 7.4 (right), I did not observe such change for average out-group mentioning percentage, with no significant difference between the two groups ($t(40) = 0.03, p = 0.97$), and qualitatively we saw a decreasing trend in the later period for the anti-Snowden group. The results suggest that, compared to the pro-Snowden group, members of the anti group engaged in increasing inter-personal interactions over time, which also increasingly focused on like-minded others.

To further validate that the minority group had a higher "internalizing" tendency in their interpersonal interactions, I compare the change of in-group mentioning network of the two groups. I construct a directed and unweighted network by drawing an edge from a user to another if the former ever mentioned the latter — this would exclude the possibility that the increasing in-group mentions was merely due to increasing interactions between some pairs. For both groups, I compare the in-group mentioning networks for the first half period (1-21 week) and second-half period. In Table 7.4, I present the key network topology metrics.

For the anti-Snowden group, comparing the later to the earlier period, I found increases in the average degree, connectedness, reciprocity and clustering coefficient, and decreases in distance, all suggesting a more tightly interconnected in-group mentioning network over time. In contrast, for the pro-Snowden group, while the size of the mentioning network decreased to

	Anti - 1	Anti - 2	Pro - 1	Pro - 2
# Nodes	**766**	566	**10539**	6260
Avg Degree	1.75	**1.86**	**2.52**	2.26
Avg Distance	**4.73**	4.18	6.10	**6.56**
Connectedness	0.11	**0.19**	**0.18**	0.12
Reciprocity	0.115	**0.118**	**0.043**	0.037
Clustering Coefficient	0.067	**0.081**	**0.062**	0.059

Table 7.4: **Network topology metrics for in-group mentioning networks over time (1st half vs. 2nd half), with larger values in bold.**

59.4% in the latter half, compared to 73.9% for the anti-Snowden group, the average distance increased, while all the other metrics decreased. Consistent with the previous data, this suggests that the in-group mentioning network of anti-Snowden group became better connected over time, but that of the pro-Snowden group became less so. Moreover, I found that the anti-Snowden mentioning network had much higher reciprocity, suggesting more reciprocal interactions, possibly more bi-directional conversations.

The above results suggest that the increasing activeness and expressiveness of anti-Snowden group could potentially be attributed to members becoming increasingly active in interacting with like-minded others. This is consistent with the "in-group bias hypothesis" that the minority group members, compared to majority group, would have a stronger tendency to seek reinforcement from other minority members, and more discriminative against out-group members,. To further verify these differences in in-group bias, I conduct content analysis of the in-group and out-group interactions

7.3.4   Content of In-Group and Out-Group Interactions

To further unpack the in-group and out-group interactions, we conducted qualitative analysis on the content of tweets containing in-group and out-group mentioning. For each of the four types of mentioning, anti $\rightarrow$ anti, anti $\rightarrow$ pro, pro $\rightarrow$ pro, pro $\rightarrow$ anti, we draw a random sample of 250 tweets from the pool. Researchers performed an iterative coding process [194] to identify the *intention* of the mentioning. Four themes of mentioning intention emerged in the codes:

1) **Conversing**. It is typically part of a dialogue or to start a conversation. Consistent with previous findings [195, 66], we observed that a significant portion of in-group conversations were to show support for like-minded others, but out-group ones mainly expressed opposition, disapproval, or to question and provoke different-minded others, e.g.:

"@mygirls3333 EXACTLY! Rachel is FULL-OF-SHIT on that! She doesn't mention #Snowden since he ran to Russia!"

"@thetomtatum Actually, that is incorrect. The majority of the public supports #Snowden."

2) **Directing**. We observed an important use of mentioning is to direct information or facts to the targeted users [1]. These tweets were often to share factual updates about the Snowden issue or relevant external resources. Examples include:

"ICYMI (07/13) Alleged #Snowden Statement Clouded With Skepticism `http://t.co/k4VdrBeZ6S` cc @LibertyLynx @20committee @catfitz"

3) **Referencing**. This consists of tweets where users cited what the mentioned users have said. Different from directly using the retweet button, we observed that they often rewrote or adapted the original tweets by shortening or summarizing them, and also, more than half of them added additional comments. They were often used to call out opinion-different users to question, criticize, or mock their statements, or to support or endorse opinion-similar users'. For example:

"@YourAnonNews: Venezuela says it will shelter #Snowden: `http://t.co/DhYb8F9UY9` #WeStandWithEdwardSnowden"

While this can be seen as a form of sharing similar to retweeting, previous research [186] documented the conversational and relational aspect of adapting and commenting through manually sharing, with the intention to publicly agree or disagree, to start a conversation, to make visible one's presence as a listener, to signal friendship or loyalty, and so forth. Given the currently wider use of the retweet button, these motivations to manually reference could be especially noteworthy.

---

[1]13.8% of "directing" tweets overlapped with the "conversing" category, as they happened in the middle of conversation

4) **Pointing**. Lastly, we found that mentioning is also used to point to a person when publicly addressing him or her. They are intended for the general audience but also making the mentioned person aware. For example:

"@LouiseMensch @grantshapps @MailOnline all gone quiet on #Snowden since Merkel concern. Thank god 4 the @guardian and #snowden for exposure"

	anti-in	anti-out	pro-in	pro-out
Conversing	50%	56%	26%	59%
Directing	28%	10%	18%	14%
Referencing	20%	17%	42%	19%
Pointing	5%	18%	15%	9%

Table 7.5: **Frequency of code occurrence in different mentioning tweets**

I present the frequency of code occurrence for each category of mentioning in Table 7.5. The following conclusions could be drawn from the results:

Echo chamber and audiencing effects: The minority group were significantly more likely to engage in in-group conversations than the majority group ($\chi^2 = 29.67, p < 0.001$). This provides explanation for the previous observation that the mentioning network of the anti-group had higher reciprocity and became more inter-connected over time, as they were more likely to form meaningful social relationships through bi-directional conversations. This also suggests that these in-group conversations could have contributed to the increasing "audiencing" practices we observed earlier. Importantly, we consider it as evidence that the minority group were more likely to engage in exchanging reinforcing opinions, and thus exhibiting stronger "echo chamber effect".

Cooperation effects: Directing was generally more likely to happen in in-group than out-group communications, and the in-group directing happened more in the anti- than the pro-Snowden group ($\chi^2 = 7.10, p < 0.01$). Qualitatively, in-group directing could be considered an cooperative behavior for people to share information or external resources that support the group position. A typical example of the anti-Snowden group directing supportive fact is:

"Americans support #Snowden? In latest poll 60% say he hurt US secu-

rity; 52% want him charged. Some support. `http://t.co/YI3Ptv4tDk`
@Skipease"

This suggests that the minority group adopted more cooperative strategy by sharing supportive resources with the alliances.

Gatewatching effects: Referencing is generally more likely to happen in-group than out-group, and pro-Snowden group had significantly more in-group referencing than that of the anti group ($\chi^2 = 27.07, p < 0.001$). It suggests that the pro-Snowden group engaged more in sharing and publicly supporting other group members to signal their group affiliation. This is again consistent with the previous observation that the majority group focused more on the "gatewatching" practice.

Taken together, the above analyses confirmed the stronger in-group bias of the numerical minority group, as compared to the majority group. The in-group bias is not only reflected in the network topology, but also and fundamentally in a stronger tendency to actively seek and extend support to each other. While I was not able to directly study the causes of such group difference in in-group bias, the observation is largely consistent with previous social science research examining in-group bias of numerical minority groups, which has associated the phenomenon with the heightened self-threat of minority groups. Therefore, defense motivation related attributes may not only lead to increased selective exposure in individuals, and also, at the aggregate level, it could result in stronger in-group bias. This work highlights that varied in-group bias not only constructs varied social and information environments for people of different opinion groups, it can also impact user behaviors, including the difference in activeness and expressiveness between users of different opinion groups.

7.3.5 Amplification (RQ4)

In this section, I study whether and how the collective use of the retweet button could amplify the presence of certain opinion groups over others (RQ4). Because a group uses the social feature more actively, it may give advantage to the peripheral nudges for their opinion, and increase the presence and impact of their viewpoint. To do so, I introduce the concept of *amplification*

to reflect the asymmetrical results of the collective use of social features. To conceptualize amplification, I adopt a "black box" view and consider the distribution of original messages from different opinion groups as the *input distribution*, and the distribution of the total retweets each group received to be the *output distribution*. Depending on the design feature studied, the "distribution" could be defined as the total number of retweets, likes, or total ratings, etc. of the messages from different groups. I define amplification to be the ratio between the output distribution and the input distribution. It allows one to quantify the size of the potential amplification effect of the collective use of a design feature, and also, to monitor the effect over time.

I study the amplification effect of the *retweeting button*. On this platform, retweeting can play two roles in amplifying a group's viewpoint: First, it can increase the number of presence of messages, thus reaching a wider audience; and second, to serve as a peripheral nudge on users' Twitter feed by a "shared" icon with the number of retweeting. Using the meta-field of the ID of the retweeted tweet, I was able to track how many times each tweet was retweeted through the retweet button. Specifically, I calculate the amplification effect by the following formula:

$$Amplification = \frac{N_{pro}/N_{anti}}{n_{pro}/n_{anti}} \tag{7.2}$$

where n_i is the number of original messages from group i, N_i is the total size after retweeting, calculated by $N_i = \sum_{j=1}^{n_i}(1 + N_{RT}(j))$, where $N_{RT}(j)$ is the number of times a tweet j got retweeted. If the amplification index is equal to 1, it means there is no amplification. If it is above (below) 1, there is an amplification effect favoring the pro (anti) group.

Figure 7.5 shows the amplification effect by week. Two conclusions can be drawn: 1) Perhaps not surprisingly, there was consistent amplification favoring the pro-Snowden (majority) group, as indicated by the amplification index being above 1 almost all the time. 2) More interestingly, the amplification effect increased over time (positive coefficient significant at $t(80) = 4.05, p < 0.001$). In the later period, the index often reached more than 2, which means that *the collective use of the retweet button could have amplified the voice of the pro-Snowden group* by more than twice of the distribution of the original messages.

133
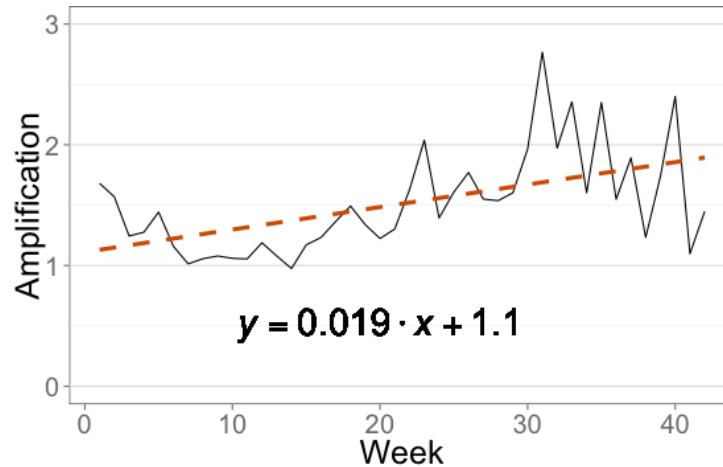
Figure 7.5: **Amplification effect of collective use of retweeting button by week**

Here I can make some inferences about the causes of the increasing amplification effect. As the anti group shifted towards more interpersonal communications, they became more disadvantaged for amplification. Not only because group members might have focused their energy on making original statements, but this type of commentary and conversational tweets are naturally less appealing for retweeting [185]. In contrast, for the pro-Snowden group, the kind of "gatewatching" practice is by nature an "amplifying" practice—by having a concentrated group of people producing original messages (low input), and having a large number of people retweeting these messages (high output).

While there could be many reasons for a group to engage in gatewatching practices, I attribute a key one to the existence of "super influencers" in the pro group, who were the target of "watching". They include well-known persons or organizations that are highly involved in the Snowden issue such as *wikileaks*, *Jesselyn Radack*, and *Jacob Appelbaum*, and dedicated activist accounts such as *free_snowden* and *NO2NSA*. In contrast, opinion leaders in the anti group are far less known, with a few political scholars, e.g. *Louise Mensch*, getting the most retweets. On the one hand, these super influencers are highly visible to attract retweeting by having a large number of followers and names that raise issue-specific attention. On the other hand, similar to what was observed in previous studies of activists, politicians and journalists on Twitter [196, 197, 174, 198, 199], I observed that many of them engaged

134

in strategically tweeting behaviors that are more likely to prompt sharing, e.g., creating or citing external resources, calling attention for support, using political and philosophical slogans, framing issues, using humor, and so forth. Examples include:

> "Edward Snowden should seek asylum in the only place truly beyond the reach of US law enforcement. Wall Street. #snowden" (wikileaks)

> "#Snowden should win the Nobel Peace Prize. Dear @Nobelprize_org, please listen to the Internet We want Snowden! He's a hero! Retweet = $\sqrt{}$."(KimDotcom)

> "My artcl: How Obama misled public when he said protected legal channels exist that #Snowden could've used http://t.co/CuGqsDoImU via @Salon"(JesselynRadack)

I quantitatively compare the activities of the "top influencers" and a group's retweeting of them between the two groups (the method to identify them was discussed in the "user classification" section). I found that, indeed, the top influencers in the pro group were far more likely to be retweeted, with an average retweeting rate of 36.44 per tweet, compared to 4.20 for the top influencers in the anti group. By analyzing the temporal changes of the percentage of tweets contributed by the top influencers (Figure 7.6 (left)),Ie found that the contributions of pro-Snowden opinion leaders were steady, and slightly increased over time, in contrast with the decreasing contribution of the anti-group opinion leaders (the difference is significant at $t(80) = 3.28, p = 0.002$). Meanwhile, there was an increasing tendency for the pro group to focus on retweeting the top influencers, as reflected by the increasing proportion of retweets directed to the top influencers (Figure 7.6 (right), the difference is significant at $t(80) = 1.92, p = 0.05$). Together, it suggests that the increasing amplification could be at least partially attributed to the increasing practices of gatewatching for the active and influential opinion leaders in the pro-Snowden group.

To summarize, the collective use of *the retweet button amplified the voices of the majority group*, and the *amplification increased over time*. This indicates that, while the production of original tweets dropped faster for the pro-Snowden group, their retweeting activities were better sustained, suggesting that the pro-group members shifted their attention from making
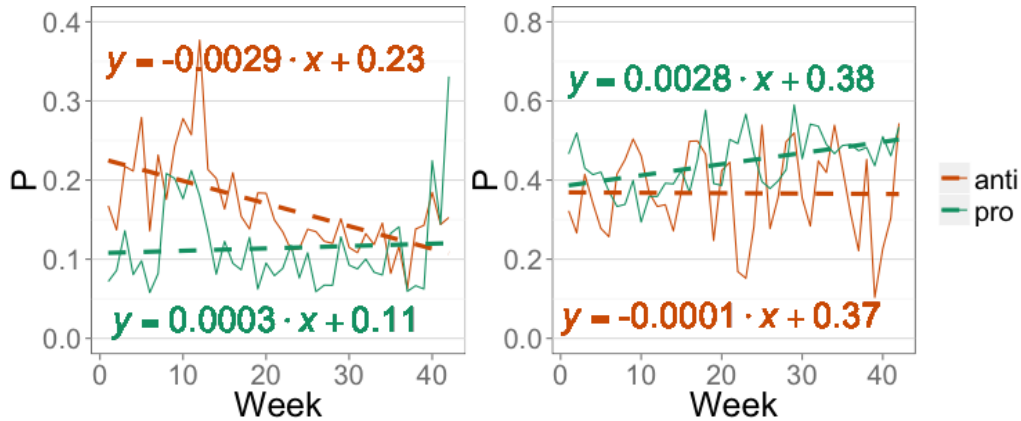
Figure 7.6: **(left) Percentage of tweets from top influencers (right) Percentage of retweets directed to top influencers' tweets**

original comments to retweeting existing tweets, especially retweeting those from opinion leaders. There are several implications from the asymmetrical results of the collective retweeting activities between the two opinion groups. One is that the peripheral nudges to the views of the majority group is much more prominent. For example, a Twitter user who is interested in checking the Snowden discussion, upon searching for the topic, would see a list of tweets. The user's attention would be drawn to those popular tweets that come with the peripheral nudge "shared by many others", which would be primarily pro-Snowden tweets. This problem would be further exacerbated by the "Top" feature of Twitter—recommendation of top tweets on the topic by taking the number of sharing into consideration. Such a scenario creates the very opposite outcome for the goal of enhancing users' exposure to diverse views.

This also means that even though the minority group increased its presence in creating original messages, its presence on the platform could still be weakened. Considering someone conducting data analysis by pooling both original and retweeted contents, he or she might reach the conclusion that the social opinions have turned towards pro-Snowden. This sends a critical warning against pooling social media data without disentangling the distorting effect of the collective use of social features, which could lead to incorrect conclusions about social opinions.

7.4 Discussion

This work sends two warnings for both designing and studying social technologies that host diverse viewpoints. The first warning is that opinion groups may differ in their selective exposure tendency, leading to some groups better inter-connected and more segregated form others. This not only limits the exposure to diverse views for the group members, but also, creates varied social environments that shape user behaviors. Specifically, I found evidence that, in the long run, topical discussions may be better sustained in opinion groups that show stronger selective exposure, and one reason is that they would be more actively engage in supportive and cooperative inter-personal interactions with like-minded others. Due to the nature of these interactions, they also tend to produce content with more expressive characteristics.

This study underlines the broader problem of studying social opinions with "pooled data" from social media without considering the behavioral differences and changes of opinion groups, and thus biases of their presence and differences in the content characteristics they create [200, 175]. Although many factors may potentially contribute to the group behavior differences, this study suggests that one factor to consider is the varied in-group bias—not just narrowly in the form of connecting with similar minded others, but also with a stronger tendency to seek confirmation from and show support to each other.

Biases introduced for social media data analytics is only one problem that group differences in in-group bias may result. Other concerns include the asymmetrical social networks that can create unequal advantages in many aspects, such as member engagement, group visibility, and information dissemination. This can have profound implication on the political landscape. For example, reports (Pew Internet Report, 2008) show that during the 2008 election, the Obama campaign enjoyed a clear advantage in terms of online voter engagement than did McCain's, with more discussions and web traffic on popular social media outlets. Many believe that such advantage played a role in Obama's successful election.

These possibilities highlight the importance of more ethical design of social media by accounting for the potential variance in in-group bias between different opinion groups. To understand, even predict, group differences in in-

group bias, a possibility is to resort to theories on the moderators of selective exposure. The accuracy-defense motivation framework could potentially be a useful tool for the purpose. For example, groups with attributes associated with defense motivation such as being a minority group, a new group being challenged by the establishment, could potentially exhibit stronger in-group bias. On the other hand, groups with attributes that are likely associated with accuracy motivation, such as with high topical interests, may potentially engage in more cross-group interactions. Future research may explore the association between the known moderators of selective exposure and group behavioral differences on social media platform, and use that knowledge to inform studies using social media data.

A second warning this study presents is that group difference in collective use of social features can asymmetrically amplify the impact of groups that use these features more actively. The amplification can happen in multiple forms. First, social sharing features can asymmetrically increase the presence of messages from different opinion groups. It may not only impact the visibility of different opinions for users of the platform, but also bias the presence of them in the data. This study warns that, when analyzing social media data, it is important to consider whether, and how to account for effects enabled by collective use of design features (e.g., whether to include retweets, and, if so, whether to give retweets the same weight as the original tweets). These are practical questions that can lead to different conclusions.

Moreover, as Study 3 (Chapter 5) shows, social features can work as peripheral nudges to increase users' attendance, favorable perception and agreement with the information. The asymmetric result of the collective use of social features thus may create asymmetric peripheral nudges towards different opinion groups, an outcome that is detrimental to the goal of enhancing diversity and facilitating public deliberation. Future research may empirically examine these implications.

Finally, I want to point out the increasing presence of opinion minorities (w.r.t. original messages) is not necessarily a rejection of the idea of "spiral of silence". Given its prediction that the entry barrier to start expressing one's opinion would be higher for opinion minorities, it implies that, on average, the minority group who are in the discussions might have higher defense motivation to begin with. With lower entry barriers for the majority

opinion group, some with lower issue engagement might be willing to join the discussion in the beginning but dropped their activities over time, or shifted to less demanding activities such as retweeting.

Chapter 8

SUMMARY AND FUTURE DIRECTIONS

The central objective of this dissertation is to understand individual and contextual factors that moderate people's propensity to engage in selective exposure in interactions with information and social technologies, and to inform design of technologies that nudge people to seek attitude-challenging information by tailoring designs based on these moderators. This research is inspired by half a century of social science research on selective exposure theories, and the emerging HCI research on designing diversity-enhancing technologies. This dissertation seek to bridge the two fields. Specifically, I leverage a theoretical framework on the moderators of selective exposure— the *accuracy-defense motivation framework*, and explore its applications in four aspects of diversity-enhancing technology designs. In the following sections, I will summarize the conclusions from this dissertation, and suggest future directions, for each of the four aspects. In the last section, I will reflect on the choice of the theoretical framework.

8.1 Designing with Positive and Negative Moderators of Selective Exposure

The first aspect of leveraging moderators of selective exposure in designing diversity-enhancing technologies is a straightforward one—to eliminate design features associated with the positive moderators of (increase) selective exposure, and adopt ones that contribute to the negative moderators of (decrease) selective exposure. As a demonstration of the idea, I showed that design features that heighten threat, and thus individuals' defense motivation, can lead to pronounced selective exposure (Study 1 in Chapter 3). Noting that sometimes these features are unintentionally created, such as

a decorative image, I emphasize the importance of understanding and inspecting potential moderators of selective exposure to avoid designs that can provoke the tendency to engage in undesired information selectivity.

Many design ideas can arise from the knowledge about moderators of selective exposure. For example, based on the recommendation of reducing defense motivation, easy access to supportive information has been robustly proved to ease cognitive dissonance and promote exposure to diverse views. This can be easily implemented with digital technologies. Other ideas, as inspired by the meta-analysis of Hart et al., include avoiding forcing users to make committed choices, avoiding challenging users' core value, and providing ways to boost users' self-confidence. Designs can also focus on promoting accuracy motivation. Balancer [104] is an example on point by priming accuracy motivation as social norms. Another general strategy is to highlight the utility value of attitude-inconsistent information.

As a general direction, future research can leverage the multi-modality and cue richness of digital technologies to carefully construct an environment that reduces cognitive dissonance and fosters open-mindedness. This may range from leveraging graphic cues to nudge users' information access path, building system features that change users' interaction patterns, to adopting novel interaction techniques such as the recent rise of conversational agents for information retrieval (e.g., chatbot). Agent interfaces may be well positioned for "nudging" users' information seeking path and attitude changes, with anthropomorphic designs and conversational interactions. For example, the idea of debate-bot was introduced recently, and this could be the next-generation platform for personal deliberation and social-opinion seeking.

A problematic observation from online cross-ideological interactions is that even if cross-group discussions do happen, instead of having meaningful opinion exchange and collective deliberation, they frequently appear to be defensive and offensive in nature [195]. The long-term effect is worrying, given that triggered defensive inclination may exacerbate individual's selective exposure. A challenge for future research would be to develop interventions to mitigate defense motivation and promote accuracy motivation in ongoing social interactions on social media. The intervention can be as simple as the automatic identification of offensive contents in out-going messages and

suggestion for revisions, to more sophisticated techniques that augment the online discussion threads.

8.2 Identification of User Groups and Use Contexts with Varied Selective Exposure Tendencies

Second, I showed that the theory on moderators of selective exposure can inform the identification of user groups and use contexts with varied levels of selective exposure tendency. By studying user interactions with an information aggregator for various topics, I showed that user group with high accuracy motivation related attributes, i.e., high topic involvement, tend to seek information in a more balanced fashion (Study 1 in Chapter 3). By exploring selective exposure in consumer medical opinion-seeking using a comment aggregator, I demonstrated that in contexts of heightened defense motivation—threat and irreversibility as in high-risk health decisions—users are more prone to selective exposure bias (Study 4 in Chapter 6). By studying group discussions with regard to a controversial topic on social media, I found evidence validating the conclusion from (offline) social science research that user group perceiving self-threat are more likely to engage in selective exposure (Study 5 in Chapter 7).

Identification of user groups with varied diversity preference serves as the first step towards personalizing diversity-enhancing designs. First of all, it helps identify user groups and use contexts that are more vulnerable to selective exposure, so that targeted interventions should be delivered. The theory also provides explanation for the observed user differences in diversity preferences with the underlying motivational mechanisms, and thus informing how to personalize designs to satisfy the needs and preferences of the different user groups and use contexts.

Moreover, identifying individual and contextual differences in diversity preference may also inform how to tailor the information retrieval and presentation with the goal of ensuring certain levels of diversity exposure without significantly impairing user favorability of the system. I propose that a more progressive "nudging" approach would be to start with the diversity level

with which the user is comfortable, and gradually increase the diversity as users' accuracy motivation increases, and/or their defense motivation drops.

Based on this idea, a future research direction would be to develop a comprehensive list of individual and contextual factors that can predict users' diversity preferences when using specific information and social technologies. While the accuracy-defense motivation framework is a helpful tool, we can also look at broader scope of research on moderators of selective exposure. Some of the factors could be context related, and seen as the "base rate" of selective exposure in the particular information-seeking context. For example, using a similar interface, I showed that in health information seeking, users are generally less likely to engage in selective exposure, compared to political information seeking tasks, possibly due to the generally low defensive inclination in health information seeking. Other factors may be related to user profiles. In addition to topical involvement, recent studies examined the moderating effect of demographic features [74] and personality traits such as close-mindedness, rebelliousness and hostility [201] on selective exposure. Other factors may arise from the interactions with the system—e.g., what information is consumed, whether threatening or confirming message is encountered, what is presented on the interface, and so forth. Such events can be easily inferred by logging users' interactions with the system.

By identifying the list of factors, we can build user models that are able to dynamically predict users' diversity preferences, and with that, develop intelligent systems that can adapt its information retrieval and presentation as the user interacts with it. Such systems would know when and how to deliver the most effective attitude-challenging information that would also be welcomed by the user. While this is an ambitious proposal, recent development in machine learning techniques for user-profiling is promising to explore for the purpose of identifying a large number of predictive features for selective exposure tendency. The challenges, however, may lie in the attainment of training data—how to reliably identify the user's diversity preference at a particular moment. Self-report is often seen as flawed, because users either cannot assess correctly or are drawn to social desirability. Large scale log data from deliberation platforms may be explored for such purposes, by inferring diversity preferences from users' actual information selectivity, using metrics similar to the Selective Exposure Index used in the above experiments.

8.3 Personalizing Nudging Designs

The third aspect of leveraging moderators of selective exposure is to tailor the diversity-enhancing nudging design based on individual and contextual differences. Because the theories of moderator provide explanation for the underlying mechanism that drives individual differences in seeking attitude-relevant information, it may inform nudging designs that best accommodate different user groups or use contexts.

In this dissertation, I made association between accuracy motivation and the dual-processing persuasion theories, which provide a theoretical account for tailoring behavior-change interventions based on motivational level, including accuracy motivation. Based on the theory, I proposed that there can be two types of nudging design for exposure to different views— central nudges that reflect the information content characteristics and facilitate deliberative choice of information, and peripheral nudges that induce positive heuristics with the attitude-challenging information, which change information behavior in a more automatic fashion. I also propose that central nudges would be more effective in reducing selective exposure for users with high accuracy motivation, and the difference of effectiveness between the two groups would be lower for peripheral nudges (Study 2 in Chapter 4).

Using a discussion platform prototype for controversial political topics, I studied the effect of a central nudge—interface cues that reflect the fine-grained positions of information content, and effect of a peripheral nudge— interface cues that are associated with expertise heuristics. I validated the proposed design guideline by demonstrating that the central nudge was only effective in reducing selective exposure for those with high accuracy motivation (Study 2 in Chapter 4). The peripheral nudge, however, worked for both groups in changing their selective exposure (Study 3 in Chapter 5). In another study, I explored another central nudge—interface cues that indicate on which issue aspect the information focuses on, in the context of seeking social opinions for medical decisions. Again, I found consistent evidence that the central nudge is more effective in changing selective exposure tendency in the context of high-risk decision than that of low-risk decision, as the former naturally constituted a high-accuracy motivation context (Study 4 in Chapter 6).

To the best of my knowledge, this series of studies is the first attempt to explore personalizing nudging designs for diversity-enhancing technologies. Recognizing that we need to move beyond the "one-size-fits-all" approach to nudge for exposure to diverse views, there are many directions that future research can explore. While there are numerous individual and contextual factors that are potentially worth tailoring for, I argue that a theory-based approach as I used would be more efficient by providing a general umbrella construct that informs tailoring strategies for many related factors. For example, accuracy motivation can be seen as a high-level attribute that encompasses learning motivation, outcome importance, and information utility goal, all of which the design guideline can be potentially applied to.

Following the same theoretical base, an immediate suggestion made by ELM would be to tailor nudging designs for ability related attributes, in that persuasion through the central route would be more likely for those with high, than low, ability. In this context, ability can also encompass many factors such as cognitive ability, general literacy, domain knowledge, and future research is needed to examine their mediating effect on different nudging designs. In some of the studies, I also noted that domain knowledge often correlates with topic involvement—thus accuracy motivation—so future studies may need to separate the effect of the two. Another immediate question would be how to tailor nudging design for varied levels of defense motivation. ELM theories suggest that as another type of motivational factor, defense motivation would also promote central processing, but in an even more biased way. So a possibility worth exploring is to conversely tailor nudging design for defense motivation, e.g., by emphasizing attitude-irrelevant features for those with high defense motivation.

Instead of using a theory-based, top-down approach to inform the tailoring strategies, we may also explore data-driven, bottom-up approaches to do so. For example, we may infer the favorability of attitude-inconsistent messages for a particular user based on his or her previous interactions, by identifying what kind of attitude-inconsistent messages he or she was interested in reading. I would like to point out that, to some extent, both the theory-based and data-driven tailoring methods aim to optimize user preference for attitude-challenging information based on its attitude-irrelevant attributes, so hybrid methods that combine the two may also be explored. For exam-

ple, as my studies suggest that fine-grained positions (especially moderately different positions), less conflicting issue aspects, and social endorsement are information attributes valued by users, these features can be included in information retrieval algorithms to identify messages that are both challenging and appealing.

8.4   Opinion Group Difference in Selective Exposure

The last aspect to apply knowledge about moderators of selective exposure is to predict selective exposure tendency at the aggregate level— the in-group bias of different opinion groups on a social media. By conducting a case study of Twitter discussions on a controversial topic, I found that the numerical minority group showed stronger selective exposure tendency, i.e., in-group bias, than the numerical majority group. The result is consistent with conclusions from social science research on in-group bias, which recognizes higher self-threat—an attribute commonly associated with defense motivation—as the cause for stronger in-group bias of minority and marginalized groups. Importantly, the study also found evidence that the difference in in-group bias can potentially drive group behavior difference. By engaging in increasing inter-personal interactions with like-minded others, the minority group became increasingly more active and produced more expressive contents, which could potentially lead to their opinions being over-represented on the platform, and in the dataset (Study 5 in Chapter 7).

To the best of my knowledge, this is the first study to explore group differences in selective exposure and its implication for studying and designing social media platforms. The result is alarming, showing that group differences in in-group bias can lead to incorrect perception of opinion prevalence for users of the platform, and wrong conclusion from social media analytics that use the data. But there is more to that. An immediate topic to study is the implication of asymmetrical social network structure resulted from varied in-group bias. It is possible that when two groups are close in size and other attributes, the one with stronger in-group bias, thus tightly inter-connected network, would gain advantages in information dissemination and other collaborative activities, such as using design features that can broaden group

impact. While I was unable to study the problem in the context of majority v.s. minority groups, I showed that asymmetrical results of the collective use of design feature could amplify the presence, reachability and impact of viewpoints of the group that use the feature more actively.

Moreover, given that many algorithms use social network related features to rank or recommend items, the results would be biased without considering the asymmetrical social networks structure between different opinion groups. All of the above problems are counter-productive to the goal of preserving and promoting exposure to diverse perspectives. Future research should explore ways to mitigate the potential biased effect resulted from group difference in in-group bias. This may include delivering interventions to groups that are more prone to selective exposure, and taking the potential asymmetrical networks structure into account when designing social features and algorithms.

In addition to exploring these potential negative outcomes, more empirical studies are needed to understand to what extent the theories of moderators of selective exposure can be applied to predict opinion group difference in in-group bias. Although I found consistent evidence here that minority groups, whose members are more likely to experience threat, showed stronger in-group bias, I have to refrain from generalizing the conclusion to all social media platforms. It is clear that some minority groups would eventually die out on a social media platform. There are many factors that can prevail the effect of in-group bias, and in-group bias may require certain conditions to happen. For example, Twitter is a network-based platform where user can curate their own information feed, and it is possible that platforms with less emphasize on network, such as a shared discussion forum, could make it harder for minority groups to engage in selective exposure. Future research needs to conduct more empirical analyses to identify both the moderators that can predict group difference in in-group bias, and the prerequisites for the moderating effect to happen.

8.5 On Theories of Moderators for Selective Exposure

In the beginning of this dissertation, the accuracy-defense motivation framework seemed to be an easy choice as the foundational theory to motivate my research. It is comprehensive, by encompassing the two major camps of works on moderators of selective exposure—one that centers around the extent of cognitive dissonance (as relevant to defense motivation), and one that focuses on the moderating effect of the information-seeking goal (as relevant to accuracy motivation). While results of my studies largely supported the theoretical framework, it is not without its limitations. The main problem is that the two constructs are not mutually exclusive in their predictors. Study 4 (Chapter 6) best illustrates the problem, as in the potential co-existence of attributes predicting defense motivation (threats salience) and accuracy motivation (outcome importance).

The occasional correlation of the two constructs has been noted in the original work on the accuracy-defense motivation theory [42], which found it to largely offset the expected moderating effect of accuracy motivation. One explanation is that defense motivation can be more easily satisfied with the automatic process (i.e., selective exposure), while satisfying accuracy motivation (i.e., correcting selective exposure) requires conscious monitoring [202]. Examples of co-existing situations include facing issues that are both high in importance and threats to core values, and being presented with both high-quality pro and counter arguments.

So our understanding on the two motivational constructs and their interactive effects needs to be deepened, calling for more empirical and theoretical research in this area. We may also seek alternative theoretical paradigms. Instead of continuing down the motivational path, some advocates a cognitive perspective, by considering that processing attitude-consistent information tends to consume less cognitive resource because of the easier retrieval of relevant attitude and beliefs [203]. Another line of research theorizes through the lens of information utility. Specifically, the higher the utility of attitude-inconsistent information is, the less likely one engages in selective exposure [97]). Considering these two alternative theories, it would be interesting to explore selective exposure tendency as a cost-utility optimization problem, and attempt to place various contextual and individual moderators within

this unified framework.

BIBLIOGRAPHY

[1] R. E. Petty and J. T. Cacioppo, *The elaboration likelihood model of persuasion.* Springer, 1986.

[2] K. Ask and P. A. Granhag, "Motivational sources of confirmation bias in criminal investigations: The need for cognitive closure," *Journal of Investigative Psychology and Offender Profiling*, vol. 2, no. 1, pp. 43–63, 2005.

[3] C. R. Mynatt, M. E. Doherty, and R. D. Tweney, "Confirmation bias in a simulated research environment: An experimental study of scientific inference," *The quarterly journal of experimental psychology*, vol. 29, no. 1, pp. 85–95, 1977.

[4] J. M. Pines, "Profiles in patient safety: confirmation bias in emergency medicine," *Academic Emergency Medicine*, vol. 13, no. 1, pp. 90–94, 2006.

[5] J. Reason, "Understanding adverse events: human factors." *Quality in health care*, vol. 4, no. 2, pp. 80–89, 1995.

[6] P. Gallimore, "Confirmation bias in the valuation process: a test for corroborating evidence," *Journal of Property Research*, vol. 13, no. 4, pp. 261–273, 1996.

[7] D. M. Podell and L. C. Soodak, "Teacher efficacy and bias in special education referrals," *The Journal of educational research*, vol. 86, no. 4, pp. 247–253, 1993.

[8] R. S. Nickerson, "Confirmation bias: A ubiquitous phenomenon in many guises." *Review of general psychology*, vol. 2, no. 2, p. 175, 1998.

[9] J. S. Fishkin, *Democracy and deliberation: New directions for democratic reform.* Cambridge Univ Press, 1991, vol. 217.

[10] J. Elster, *Deliberative democracy.* Cambridge University Press, 1998, vol. 1.

[11] D. C. Mutz, *Hearing the other side: Deliberative versus participatory democracy.* Cambridge University Press, 2006.

[12] S. Benhabib, *Situating the self: Gender, community, and postmodernism in contemporary ethics.* Psychology Press, 1992.

[13] C. R. Sunstein, *Going to extremes: How like minds unite and divide.* Oxford University Press, 2009.

[14] J. Habermas, *The structural transformation of the public sphere: An inquiry into a category of bourgeois society.* MIT press, 1991.

[15] L. Festinger, *A theory of cognitive dissonance.* Stanford university press, 1962, vol. 2.

[16] C. R. Sunstein, *Republic. com 2.0.* Princeton University Press, 2009.

[17] D. C. Mutz, "Facilitating communication across lines of political difference: The role of mass media," in *American Political Science Association*, vol. 95, no. 01. Cambridge Univ Press, 2001, pp. 97–114.

[18] M. Prior, *Post-broadcast democracy: How media choice increases inequality in political involvement and polarizes elections.* Cambridge University Press, 2007.

[19] E. Pariser, *The filter bubble: How the new personalized web is changing what we read and how we think.* Penguin, 2011.

[20] T. Kriplean, J. Morgan, D. Freelon, A. Borning, and L. Bennett, "Supporting reflective public thought with considerit," in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work.* ACM, 2012, pp. 265–274.

[21] S. Faridani, E. Bitton, K. Ryokai, and K. Goldberg, "Opinion space: a scalable tool for browsing online comments," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2010, pp. 1175–1184.

[22] D. Arnott, "Cognitive biases and decision support systems development: a design science approach," *Information Systems Journal*, vol. 16, no. 1, pp. 55–78, 2006.

[23] E. P. Baumer, F. Polletta, N. Pierski, C. Celaya, K. Rosenblatt, and G. K. Gay, "Developing computational supports for frame reflection," 2013.

[24] N. Benn and A. Macintosh, "Argument visualization for eparticipation: towards a research agenda and prototype tool," in *Electronic Participation.* Springer, 2011, pp. 60–73.

[25] K. Dave, M. Wattenberg, and M. Muller, "Flash forums and forum-reader: navigating a new kind of large-scale online discussion," in *Proceedings of the 2004 ACM conference on Computer supported cooperative work*. ACM, 2004, pp. 232–241.

[26] S. Park, S. Kang, S. Chung, and J. Song, "Newscube: delivering multiple aspects of news to mitigate media bias," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 443–452.

[27] R. Awadallah, M. Ramanath, and G. Weikum, "Opinionetit: understanding the opinions-people network for politically controversial topics," in *proceedings of the 20th ACM international conference on Information and knowledge management*. ACM, 2011, pp. 2481–2484.

[28] B. Semaan, H. Faucett, S. P. Robertson, M. Maruyama, and S. Douglas, "Designing political deliberation environments to support interactions in the public sphere," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2015, pp. 3167–3176.

[29] T. Kriplean, M. Toomim, J. Morgan, A. Borning, and A. Ko, "Is this what you meant?: promoting listening on the web with reflect," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 1559–1568.

[30] J. E. Introne, "Supporting group decisions by mediating deliberation to improve information pooling," in *Proceedings of the ACM 2009 international conference on Supporting group work*. ACM, 2009, pp. 189–198.

[31] R. W. White and E. Horvitz, "Cyberchondria: studies of the escalation of medical concerns in web search," *ACM Transactions on Information Systems (TOIS)*, vol. 27, no. 4, p. 23, 2009.

[32] C. Lauckner and G. Hsieh, "The presentation of health-related search results and its impact on negative emotional outcomes," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 333–342.

[33] A. Lau and E. Coiera, "Impact of web searching and social feedback on consumer decision making: a prospective online experiment," *Journal of medical Internet research*, vol. 10, no. 1, p. e2, 2008.

[34] P. Resnick, R. K. Garrett, T. Kriplean, S. A. Munson, and N. J. Stroud, "Bursting your (filter) bubble: strategies for promoting diverse exposure," in *Proceedings of the 2013 conference on Computer supported cooperative work companion*. ACM, 2013, pp. 95–100.

[35] C. R. Sunstein and R. Thaler, "Nudge," *The politics of libertarian paternalism. New Haven*, 2008.

[36] S. A. Munson and P. Resnick, "The prevalence of political discourse in non-political blogs," in *Fifth International AAAI Conference on Weblogs and Social Media*, 2011.

[37] J. Stromer-Galley, "Diversity of political conversation on the internet: Users' perspectives," *Journal of Computer-Mediated Communication*, vol. 8, no. 3, pp. 0–0, 2003.

[38] E. Gilbert, T. Bergstrom, and K. Karahalios, "Blogs are echo chambers: Blogs are echo chambers," in *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on.* IEEE, 2009, pp. 1–10.

[39] K. Yatani, M. Novati, A. Trusty, and K. N. Truong, "Review spotlight: a user interface for summarizing user-generated reviews using adjective-noun word pairs," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2011, pp. 1541–1550.

[40] E. P. Baumer, V. Khovanskaya, M. Matthews, L. Reynolds, V. Schwanda Sosik, and G. Gay, "Reviewing reflection: on the use of reflection in interactive system design," in *Proceedings of the 2014 conference on Designing interactive systems.* ACM, 2014, pp. 93–102.

[41] D. O. Sears and J. L. Freedman, "Selective exposure to information: A critical review," *Public Opinion Quarterly*, vol. 31, no. 2, pp. 194–213, 1967.

[42] W. Hart, D. Albarracín, A. H. Eagly, I. Brechan, M. J. Lindberg, and L. Merrill, "Feeling validated versus being correct: a meta-analysis of selective exposure to information." *Psychological bulletin*, vol. 135, no. 4, p. 555, 2009.

[43] D. Frey and D. Stahlberg, "Selection of information after receiving more or less reliable self-threatening information," *Personality and Social Psychology Bulletin*, vol. 12, no. 4, pp. 434–441, 1986.

[44] D. Albarracín and A. L. Mitchell, "The role of defensive confidence in preference for proattitudinal information: How believing that one is strong can sometimes be a defensive weakness," *Personality and Social Psychology Bulletin*, vol. 30, no. 12, pp. 1565–1584, 2004.

[45] J. W. Brehm and A. R. Cohen, "Explorations in cognitive dissonance." 1962.

[46] J. M. Olson and J. Stone, "The influence of behavior on attitudes." 2005.

[47] D. Albarracin, "Cognition in persuasion: An analysis of information processing in response to persuasive communications," *Advances in experimental social psychology*, vol. 34, pp. 61–130, 2002.

[48] L. K. Canon, "Self-confidence and selective exposure to information," *Conflict, decision, and dissonance*, vol. 1, pp. 83–95, 1964.

[49] P. Fischer, A. Kastenmüller, T. Greitemeyer, J. Fischer, D. Frey, and D. Crelley, "Threat and selective exposure: The moderating role of threat and decision context on confirmatory information search after decisions." *Journal of Experimental Psychology: General*, vol. 140, no. 1, p. 51, 2011.

[50] E. Jonas, J. Greenberg, and D. Frey, "Connecting terror management and dissonance theory: Evidence that mortality salience increases the preference for supporting information after decisions," *Personality and Social Psychology Bulletin*, vol. 29, no. 9, pp. 1181–1189, 2003.

[51] N. A. Valentino, A. J. Banks, V. L. Hutchings, and A. K. Davis, "Selective exposure in the internet age: The interaction between anxiety and information utility," *Political Psychology*, vol. 30, no. 4, pp. 591–613, 2009.

[52] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 us election: divided they blog," in *Proceedings of the 3rd international workshop on Link discovery*. ACM, 2005, pp. 36–43.

[53] M. D. Conover, B. Gonçalves, A. Flammini, and F. Menczer, "Partisan asymmetries in online political activity," *EPJ Data Science*, vol. 1, no. 1, pp. 1–19, 2012.

[54] B. Simon, B. Aufderheide, and C. Kampmeier, "The social psychology of minority-majority relations," *Self and social identity*, pp. 278–297, 2004.

[55] J. Klayman, "Varieties of confirmation bias," *Psychology of learning and motivation*, vol. 32, pp. 385–418, 1995.

[56] A. H. Eagly and S. Chaiken, *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers, 1993.

[57] C. J. Nemeth, "Differential contributions of majority and minority influence." *Psychological review*, vol. 93, no. 1, p. 23, 1986.

[58] M. L. Maznevski, "Understanding our differences: Performance in decision-making groups with diverse members," *Human relations*, vol. 47, no. 5, pp. 531–552, 1994.

[59] T. Mendelberg, "The deliberative citizen: Theory and evidence," *Political decision making, deliberation and participation*, vol. 6, no. 1, pp. 151–193, 2002.

[60] S. Iyengar and K. S. Hahn, "Red media, blue media: Evidence of ideological selectivity in media use," *Journal of Communication*, vol. 59, no. 1, pp. 19–39, 2009.

[61] N. J. Stroud, "Media use and political predispositions: Revisiting the concept of selective exposure," *Political Behavior*, vol. 30, no. 3, pp. 341–366, 2008.

[62] R. Brown, "Social identity theory: Past achievements, current problems and future challenges," *European journal of social psychology*, vol. 30, no. 6, pp. 745–778, 2000.

[63] E. L. Glaeser and B. A. Ward, "Myths and realities of american political geography," National Bureau of Economic Research, Tech. Rep., 2005.

[64] D. S. Massey, J. Rothwell, and T. Domina, "The changing bases of segregation in the united states," *The Annals of the American Academy of Political and Social Science*, vol. 626, no. 1, pp. 74–90, 2009.

[65] E. Hargittai, J. Gallo, and M. Kane, "Cross-ideological discussions among conservative and liberal bloggers," *Public Choice*, vol. 134, no. 1-2, pp. 67–86, 2008.

[66] S. Yardi and D. Boyd, "Dynamic debates: An analysis of group polarization over time on twitter," *Bulletin of Science, Technology & Society*, vol. 30, no. 5, pp. 316–327, 2010.

[67] M. Gentzkow and J. M. Shapiro, "Ideological segregation online and offline," National Bureau of Economic Research, Tech. Rep., 2010.

[68] N. Negroponte, *Being digital.* Vintage, 1996.

[69] S. Flaxman, S. Goel, and J. M. Rao, "Ideological segregation and the effects of social media on news consumption," *Available at SSRN 2363701*, 2013.

[70] J. Kelly, M. Smith, and D. Fisher, "Opinion diversity in online political discussion networks," *Online Deliberation 2005*, 2005.

[71] M. Conover, J. Ratkiewicz, M. R. Francisco, B. Gonçalves, F. Menczer, and A. Flammini, "Political polarization on twitter." *ICWSM*, vol. 133, pp. 89–96, 2011.

[72] R. K. Garrett, "Politically motivated reinforcement seeking: Reframing the selective exposure debate," *Journal of Communication*, vol. 59, no. 4, pp. 676–699, 2009.

[73] S. A. Munson and P. Resnick, "Presenting diverse political opinions: how and how much," in *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2010, pp. 1457–1466.

[74] S. Knobloch-Westerwick and J. Meng, "Looking the other way selective exposure to attitude-consistent and counterattitudinal political information," *Communication Research*, vol. 36, no. 3, pp. 426–448, 2009.

[75] J. Mills, E. Aronson, and H. Robinson, "Selectivity in exposure to information." *The Journal of Abnormal and Social Psychology*, vol. 59, no. 2, p. 250, 1959.

[76] A. Lev-On and B. Manin, "Happy accidents: Deliberation and on-line exposure to opposing views," *Online deliberation: Design, research and practice*, pp. 105–122, 2009.

[77] R. K. Garrett and P. Resnick, "Resisting political fragmentation on the internet," *Daedalus*, vol. 140, no. 4, pp. 108–120, 2011.

[78] A. L. Holbrook, M. K. Berent, J. A. Krosnick, P. S. Visser, and D. S. Boninger, "Attitude importance and the accumulation of attitude-relevant knowledge in memory." *Journal of personality and social psychology*, vol. 88, no. 5, p. 749, 2005.

[79] D. Frey, "Recent research on selective exposure to information," *Advances in experimental social psychology*, vol. 19, no. 1, pp. 41–80, 1986.

[80] R. Cohen and D. Ruths, "Classifying political orientation on twitter: It's not easy!" in *ICWSM*, 2013.

[81] D. Frey and M. Irle, "Some conditions to produce a dissonance and an incentive effect in a forced-compliancesituationx," *European Journal of Social Psychology*, vol. 2, no. 1, pp. 45–54, 1972.

[82] P. D. Sweeney and K. L. Gruber, "Selective exposure: Voter information preferences and the watergate affair." *Journal of Personality and Social Psychology*, vol. 46, no. 6, p. 1208, 1984.

[83] N. Schwarz, D. Frey, and M. Kumpf, "Interactive effects of writing and reading a persuasive essay on attitude change and selective exposure," *Journal of Experimental Social Psychology*, vol. 16, no. 1, pp. 1–17, 1980.

[84] J. L. Freedman, "Confidence, utility, and selective exposure: a partial replication." *Journal of personality and social psychology*, vol. 2, no. 5, p. 778, 1965.

[85] A. Lowin, "Further evidence for an approach-avoidance interpretation of selective exposure," *Journal of Experimental Social Psychology*, vol. 5, no. 3, pp. 265–271, 1969.

[86] D. Frey, "Reversible and irreversible decisions preference for consonant information as a function of attractiveness of decision alternatives," *Personality and Social Psychology Bulletin*, vol. 7, no. 4, pp. 621–626, 1981.

[87] E. Jonas, V. Graupmann, and D. Frey, "The influence of mood on the search for supporting versus conflicting information: Dissonance reduction as a means of mood regulation?" *Personality and Social Psychology Bulletin*, vol. 32, no. 1, pp. 3–15, 2006.

[88] L. A. Brannon, M. J. Tagler, and A. H. Eagly, "The moderating role of attitude strength in selective exposure to information," *Journal of Experimental Social Psychology*, vol. 43, no. 4, pp. 611–617, 2007.

[89] P. Fischer, E. Jonas, D. Frey, and A. Kastenmüller, "Selective exposure and decision framing: The impact of gain and loss framing on confirmatory information search after decisions," *Journal of Experimental Social Psychology*, vol. 44, no. 2, pp. 312–320, 2008.

[90] D. Zillmann, L. Chen, S. Knobloch, and C. Callison, "Effects of lead framing on selective exposure to internet news reports," *Communication Research*, vol. 31, no. 1, pp. 58–81, 2004.

[91] W. R. Neuman, *The future of the mass audience.* Cambridge University Press, 1991.

[92] S. Chaiken, W. Wood, and A. H. Eagly, "Principles of persuasion." 1996.

[93] B. T. Johnson and A. H. Eagly, "Effects of involvement on persuasion: A meta-analysis." *Psychological Bulletin*, vol. 106, no. 2, p. 290, 1989.

[94] S. L. Sargent, "Image effects on selective exposure to computer-mediated news stories," *Computers in Human Behavior*, vol. 23, no. 1, pp. 705–726, 2007.

[95] S. M. Smith, L. R. Fabrigar, D. M. Powell, and M.-J. Estrada, "The role of information-processing capacity and goals in attitude-congruent selective exposure effects," *Personality and Social Psychology Bulletin*, 2007.

[96] S. Chen and S. Chaiken, "The heuristic-systematic model in its broader context," *Dual-process theories in social psychology*, pp. 73–96, 1999.

[97] C. K. Atkin, "Informational utility and selective exposure to entertainment media," *Selective exposure to communication*, pp. 63–91, 1985.

[98] C. S. Park, "Does twitter motivate involvement in politics? tweeting, opinion leadership, and political engagement," *Computers in Human Behavior*, vol. 29, no. 4, pp. 1641–1648, 2013.

[99] A. H. Oh, H.-J. Lee, and Y.-M. Kim, "User evaluation of a system for classifying and displaying political viewpoints of weblogs." in *ICWSM*, 2009.

[100] Y. Jiang, Q. V. Liao, Q. Cheng, R. B. Berlin, and B. R. Schatz, "Designing and evaluating a clustering system for organizing and integrating patient drug outcomes in personal health messages," in *AMIA Annual Symposium Proceedings*, vol. 2012. American Medical Informatics Association, 2012, p. 417.

[101] E. P. Baumer, C. Cipriani, M. Davis, G. He, J. Kang, J. Jeffrey-Wilensky, J. Lee, J. Zupnick, and G. K. Gay, "Broadening exposure, questioning opinions, and reading patterns with reflext: a computational support for frame reflection," *Journal of Information Technology & Politics*, vol. 11, no. 1, pp. 45–63, 2014.

[102] E. Graells-Garrido, M. Lalmas, and R. Baeza-Yates, "Data portraits and intermediary topics: Encouraging exploration of politically diverse profiles," *arXiv preprint arXiv:1601.00481*, 2016.

[103] C. Schwind and J. Buder, "Reducing confirmation bias and evaluation bias: When are preference-inconsistent recommendations effective–and when not?" *Computers in Human Behavior*, vol. 28, no. 6, pp. 2280–2290, 2012.

[104] S. A. Munson, S. Y. Lee, and P. Resnick, "Encouraging reading of diverse political viewpoints with a browser widget." in *ICWSM*, 2013.

[105] S. Messing and S. J. Westwood, "Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online," *Communication Research*, p. 0093650212466406, 2012.

[106] E. A. S. A. Munson, "Social cues and interest in reading political news stories," 2015.

[107] B. H. Bornstein and A. C. Emler, "Rationality in medical decision making: a review of the literature on doctors decision-making biases," *Journal of evaluation in clinical practice*, vol. 7, no. 2, pp. 97–107, 2001.

[108] G. M. Leydon, M. Boulton, C. Moynihan, A. Jones, J. Mossman, M. Boudioni, and K. McPherson, "Cancer patients' information needs and information seeking behaviour: in depth interview study," *Bmj*, vol. 320, no. 7239, pp. 909–913, 2000.

[109] D. E. Brashers, D. J. Goldsmith, and E. Hsieh, "Information seeking and avoiding in health contexts," *Human Communication Research*, vol. 28, no. 2, pp. 258–271, 2002.

[110] M. Evans, A. Shaw, E. A. Thompson, S. Falk, P. Turton, T. Thompson, and D. Sharp, "Decisions to use complementary and alternative medicine (cam) by male cancer patients: information-seeking roles and types of evidence used," *BMC Complementary and Alternative Medicine*, vol. 7, no. 1, p. 25, 2007.

[111] K. Noguchi, D. Albarracín, M. R. Durantini, and L. R. Glasman, "Who participates in which health promotion programs? a meta-analysis of motivations underlying enrollment and retention in hiv-prevention interventions." *Psychological Bulletin*, vol. 133, no. 6, p. 955, 2007.

[112] S. Chancellor, J. A. Pater, T. Clear, E. Gilbert, and M. De Choudhury, "# thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities," in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing.* ACM, 2016, pp. 1201–1213.

[113] A. Y. Lau and E. W. Coiera, "Do people experience cognitive biases while searching for information?" *Journal of the American Medical Informatics Association*, vol. 14, no. 5, pp. 599–608, 2007.

[114] A. Keselman, A. C. Browne, and D. R. Kaufman, "Consumer health information seeking as hypothesis testing," *Journal of the American Medical Informatics Association*, vol. 15, no. 4, pp. 484–495, 2008.

[115] R. White, "Beliefs and biases in web search," in *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval.* ACM, 2013, pp. 3–12.

[116] C. Christensen, "Wave-riding and hashtag-jumping: Twitter, minority third parties and the 2012 us elections," *Information, Communication & Society*, vol. 16, no. 5, pp. 646–666, 2013.

[117] I. Weber, V. R. K. Garimella, and A. Batayneh, "Secular vs. islamist polarization in egypt on twitter," in *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.* ACM, 2013, pp. 290–297.

159

[118] J. Borge-Holthoefer, W. Magdy, K. Darwish, and I. Weber, "Content and network dynamics behind egyptian political polarization on twitter," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 2015, pp. 700–711.

[119] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.

[120] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury, "Twitter power: Tweets as electronic word of mouth," *Journal of the American society for information science and technology*, vol. 60, no. 11, pp. 2169–2188, 2009.

[121] S. Stieglitz and N. Krüger, "Analysis of sentiments in corporate twitter communication–a case study on an issue of toyota," *Analysis*, vol. 1, pp. 1–2011, 2011.

[122] B. Simon, B. Aufderheide, and C. Kampmeier, "The social psychology of minority-majority relations," *Blackwell handbook of social psychology: Intergroup processes*, pp. 303–323, 2001.

[123] G. J. Leonardelli and M. B. Brewer, "Minority and majority discrimination: When and why," *Journal of Experimental Social Psychology*, vol. 37, no. 6, pp. 468–485, 2001.

[124] J. C. Zhao, W.-T. Fu, H. Zhang, S. Zhao, and H. Duh, "To risk or not to risk?: Improving financial risk taking of older adults by online social information," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 2015, pp. 95–104.

[125] R. E. Petty and J. T. Cacioppo, "Involvement and persuasion: Tradition versus integration." 1990.

[126] B. Holton and T. Pyszczynski, "Biased information search in the interpersonal domain," *Personality and Social Psychology Bulletin*, vol. 15, no. 1, pp. 42–51, 1989.

[127] S. Himmelfarb, "The measurement of attitudes," *The psychology of attitudes*, pp. 23–87, 1993.

[128] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *psychometrika*, vol. 16, no. 3, pp. 297–334, 1951.

[129] M. J. Dutta-Bergman, "Complementarity in consumption of news types across traditional and new media," *Journal of Broadcasting & Electronic Media*, vol. 48, no. 1, pp. 41–60, 2004.

[130] S. Lewandowsky, U. K. Ecker, C. M. Seifert, N. Schwarz, and J. Cook, "Misinformation and its correction continued influence and successful debiasing," *Psychological Science in the Public Interest*, vol. 13, no. 3, pp. 106–131, 2012.

[131] S. Chaiken and A. H. Eagly, "Heuristic and systematic information processing within and," *Unintended thought*, vol. 212, pp. 212–252, 1989.

[132] R. M. Shiffrin and W. Schneider, "Controlled and automatic human information processing: Ii. perceptual learning, automatic attending and a general theory." *Psychological review*, vol. 84, no. 2, p. 127, 1977.

[133] A. Tversky and D. Kahneman, "The framing of decisions and the psychology of choice," *Science*, vol. 211, no. 4481, pp. 453–458, 1981.

[134] S. Chaiken, "The heuristic model of persuasion," in *Social influence: the ontario symposium*, vol. 5, 1987, pp. 3–39.

[135] C. R. Sunstein, *Why nudge?: the politics of libertarian paternalism.* Yale University Press, 2014.

[136] B. J. Fogg, "Persuasive technology: using computers to change what we think and do," *Ubiquity*, vol. 2002, no. December, p. 5, 2002.

[137] M. Kay, C. Matuszek, and S. A. Munson, "Unequal representation and gender stereotypes in image search results for occupations," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* ACM, 2015, pp. 3819–3828.

[138] R. Orji, J. Vassileva, and R. L. Mandryk, "Modeling the efficacy of persuasive strategies for different gamer types in serious games for health," *User Modeling and User-Adapted Interaction*, vol. 24, no. 5, pp. 453–498, 2014.

[139] C. M. Angst and R. Agarwal, "Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion," *MIS quarterly*, vol. 33, no. 2, pp. 339–370, 2009.

[140] S.-C. Yang, W.-C. Hung, K. Sung, and C.-K. Farn, "Investigating initial trust toward e-tailers from the elaboration likelihood model perspective," *Psychology & Marketing*, vol. 23, no. 5, pp. 429–445, 2006.

[141] T. Zhou, "Understanding users initial trust in mobile banking: An elaboration likelihood perspective," *Computers in Human Behavior*, vol. 28, no. 4, pp. 1518–1525, 2012.

[142] A. Kobsa, H. Cho, and B. P. Knijnenburg, "The effect of personalization provider characteristics on privacy attitudes and behaviors: An elaboration likelihood model approach," *Journal of the Association for Information Science and Technology*, 2016.

[143] T. Dinev and P. Hart, "An extended privacy calculus model for e-commerce transactions," *Information Systems Research*, vol. 17, no. 1, pp. 61–80, 2006.

[144] P. Pirolli and S. Card, "Information foraging." *Psychological review*, vol. 106, no. 4, p. 643, 1999.

[145] E. H. Chi, P. Pirolli, K. Chen, and J. Pitkow, "Using information scent to model user information needs and actions and the web," in *Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM, 2001, pp. 490–497.

[146] M. J. Salganik, P. S. Dodds, and D. J. Watts, "Experimental study of inequality and unpredictability in an artificial cultural market," *science*, vol. 311, no. 5762, pp. 854–856, 2006.

[147] H.-S. Kim and S. S. Sundar, "Using interface cues in online health community boards to change impressions and encourage user contribution," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2011, pp. 599–608.

[148] K. Sweeny, D. Melnyk, W. Miller, and J. A. Shepperd, "Information avoidance: Who, what, when, and why." *Review of General Psychology*, vol. 14, no. 4, p. 340, 2010.

[149] V. Vydiswaran, C. Zhai, D. Roth, and P. Pirolli, "Biastrust: Teaching biased users about controversial topics," in *Proceedings of the 21st ACM international conference on Information and knowledge management.* ACM, 2012, pp. 1905–1909.

[150] E. Aronson, J. A. Turner, and J. M. Carlsmith, "Communicator credibility and communication discrepancy as determinants of opinion change." *The Journal of Abnormal and Social Psychology*, vol. 67, no. 1, p. 31, 1963.

[151] G. Marks and N. Miller, "Ten years of research on the false-consensus effect: An empirical and theoretical review." *Psychological Bulletin*, vol. 102, no. 1, p. 72, 1987.

[152] B. Fischhoff, P. Slovic, and S. Lichtenstein, "Knowing with certainty: The appropriateness of extreme confidence." *Journal of Experimental Psychology: Human perception and performance*, vol. 3, no. 4, p. 552, 1977.

[153] J. L. Fleiss, "Measuring nominal scale agreement among many raters." *Psychological bulletin*, vol. 76, no. 5, p. 378, 1971.

[154] M. D. Mulvenna, S. S. Anand, and A. G. Büchner, "Personalization on the net using web mining: introduction," *Communications of the ACM*, vol. 43, no. 8, pp. 122–125, 2000.

[155] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and trends in information retrieval*, vol. 2, no. 1-2, pp. 1–135, 2008.

[156] D. X. Zhou, P. Resnick, and Q. Mei, "Classifying the political leaning of news articles and users from user votes." in *ICWSM*, 2011.

[157] S. Knobloch-Westerwick, N. Sharma, D. L. Hansen, and S. Alter, "Impact of popularity indications on readers' selective exposure to online news," *Journal of broadcasting & electronic media*, vol. 49, no. 3, pp. 296–313, 2005.

[158] P. Dourish and M. Chalmers, "Running out of space: Models of information navigation," in *Short paper presented at HCI*, vol. 94, 1994, pp. 23–26.

[159] W. Willett, J. Heer, and M. Agrawala, "Scented widgets: Improving navigation cues with embedded visualizations," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 13, no. 6, pp. 1129–1136, 2007.

[160] S. S. Bateman, C. A. Gutwin, and G. I. McCalla, "Social navigation for loosely-coupled information seeking in tightly-knit groups using webwear," in *Proceedings of the 2013 conference on Computer supported cooperative work.* ACM, 2013, pp. 955–966.

[161] S. Bochner and C. A. Insko, "Communicator discrepancy, source credibility, and opinion change." *Journal of Personality and Social Psychology*, vol. 4, no. 6, p. 614, 1966.

[162] R. R. Dholakia and B. Sternthal, "Highly credible sources: persuasive facilitators or persuasive liabilities?" *Journal of Consumer Research*, pp. 223–232, 1977.

[163] D. O. Case, J. E. Andrews, J. D. Johnson, and S. L. Allard, "Avoiding versus seeking: the relationship of information seeking to avoidance, blunting, coping, dissonance, and related concepts*," *Journal of the Medical Library Association*, vol. 93, no. 3, p. 353, 2005.

[164] S. M. Smith, L. R. Fabrigar, and M. E. Norris, "Reflecting on six decades of selective exposure research: Progress, challenges, and opportunities," *Social and Personality Psychology Compass*, vol. 2, no. 1, pp. 464–493, 2008.

[165] P. Fischer and T. Greitemeyer, "A new look at selective-exposure effects an integrative model," *Current Directions in Psychological Science*, vol. 19, no. 6, pp. 384–389, 2010.

[166] D. T. Wegener and R. E. Petty, "Mood management across affective states: the hedonic contingency hypothesis." *Journal of personality and social psychology*, vol. 66, no. 6, p. 1034, 1994.

[167] E. Jonas and D. Frey, "Information search and presentation in advisor–client interactions," *Organizational Behavior and Human Decision Processes*, vol. 91, no. 2, pp. 154–168, 2003.

[168] Z. Pawlak and R. Sowinski, "Rough set approach to multi-attribute decision analysis," *European Journal of Operational Research*, vol. 72, no. 3, pp. 443–459, 1994.

[169] G. R. Bergus, I. P. Levin, and A. S. Elstein, "Presenting risks and benefits to patients," *Journal of general internal medicine*, vol. 17, no. 8, pp. 612–617, 2002.

[170] J. H. Hibbard and E. Peters, "Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice," *Annual Review of Public Health*, vol. 24, no. 1, pp. 413–433, 2003.

[171] J. Mankoff, K. Kuksenok, S. Kiesler, J. A. Rode, and K. Waldman, "Competing online viewpoints and models of chronic illness," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2011, pp. 589–598.

[172] L. M. Smith, L. Zhu, K. Lerman, and Z. Kozareva, "The role of social media in the discussion of controversial topics," in *Social Computing (SocialCom), 2013 International Conference on.* IEEE, 2013, pp. 236–243.

[173] D. C. Mutz, "The consequences of cross-cutting networks for political participation," *American Journal of Political Science*, pp. 838–855, 2002.

[174] E. Mustafaraj, S. Finn, C. Whitlock, and P. T. Metaxas, "Vocal minority versus silent majority: Discovering the opionions of the long tail," in *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on.* IEEE, 2011, pp. 103–110.

[175] D. Ruths, J. Pfeffer et al., "Social media for large studies of behavior," *Science*, vol. 346, no. 6213, pp. 1063–1064, 2014.

[176] L. Mercken, M. Candel, P. Willems, and H. De Vries, "Disentangling social selection and social influence effects on adolescent smoking: the importance of reciprocity in friendships," *Addiction*, vol. 102, no. 9, pp. 1483–1492, 2007.

[177] C. R. Sunstein, "The law of group polarization," *Journal of political philosophy*, vol. 10, no. 2, pp. 175–195, 2002.

[178] E. Noelle-Neumann, "The spiral of silence a theory of public opinion," *Journal of communication*, vol. 24, no. 2, pp. 43–51, 1974.

[179] D. A. Scheufle and P. Moy, "Twenty-five years of the spiral of silence: A conceptual review and empirical outlook," *International journal of public opinion research*, vol. 12, no. 1, pp. 3–28, 2000.

[180] S. Bikhchandani, D. Hirshleifer, and I. Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, pp. 992–1026, 1992.

[181] "Post-abc poll: Nsa surveillance and edward snowden," July 2013, http://www.washingtonpost.com.

[182] "Public split over impact of nsa leak, but most want snowden prosecuted," June 2013, http://www.people-press.org.

[183] "Poll results: Snowden," March 2014, https://today.yougov.com.

[184] "How twitter reacted to the snowden interview," May 2014, http://www.nbcnews.com.

[185] B. Suh, L. Hong, P. Pirolli, and E. H. Chi, "Want to be retweeted? large scale analytics on factors impacting retweet in twitter network," in *Social computing (socialcom), 2010 ieee second international conference on.* IEEE, 2010, pp. 177–184.

[186] D. Boyd, S. Golder, and G. Lotan, "Tweet, tweet, retweet: Conversational aspects of retweeting on twitter," in *System Sciences (HICSS), 2010 43rd Hawaii International Conference on.* IEEE, 2010, pp. 1–10.

[187] P. H. Calais Guerra, A. Veloso, W. Meira Jr, and V. Almeida, "From bias to opinion: a transfer-learning approach to real-time sentiment analysis," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, 2011, pp. 150–158.

[188] H. Gao, J. Mahmud, J. Chen, J. Nichols, and M. Zhou, "Modeling user attitude toward controversial topics in online social media," in *Eighth International AAAI Conference on Weblogs and Social Media*, 2014.

[189] M. J. Welch, U. Schonfeld, D. He, and J. Cho, "Topical semantics of twitter links," in *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 2011, pp. 327–336.

[190] M. Cha, H. Haddadi, F. Benevenuto, and P. K. Gummadi, "Measuring user influence in twitter: The million follower fallacy." *ICWSM*, vol. 10, no. 10-17, p. 30, 2010.

[191] M. De Choudhury, A. Monroy-Hernandez, and G. Mark, "Narco emotions: affect and desensitization in social media during the mexican drug war," in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 2014, pp. 3563–3572.

[192] A. Bruns and S. Stieglitz, "Quantitative approaches to comparing communication patterns on twitter," *Journal of Technology in Human Services*, vol. 30, no. 3-4, pp. 160–185, 2012.

[193] A. Bruns and S. Stieglitz, "Towards more systematic twitter analysis: Metrics for tweeting activities," *International Journal of Social Research Methodology*, vol. 16, no. 2, pp. 91–108, 2013.

[194] J. A. Maxwell, *Qualitative research design: An interactive approach: An interactive approach*. Sage, 2012.

[195] Z. Liu and I. Weber, "Is twitter a public sphere for online conflicts? a cross-ideological and cross-hierarchical look," in *Social Informatics*. Springer, 2014, pp. 336–347.

[196] A. E. Holton and S. C. Lewis, "Journalists, social media, and the use of humor on twitter," *Electronic Journal of Communication*, vol. 21, no. 1/2, 2011.

[197] S. Lindgren and R. Lundström, "Pirate culture and hacktivist mobilization: The cultural and social protocols of# wikileaks on twitter," *New Media & Society*, vol. 13, no. 6, pp. 999–1018, 2011.

[198] T. Poell and E. Borra, "Twitter, youtube, and flickr as platforms of alternative journalism: The social media account of the 2010 toronto g20 protests," *Journalism*, vol. 13, no. 6, pp. 695–713, 2012.

[199] W. L. Youmans and J. C. York, "Social media and the activist toolkit: User agreements, corporate interests, and the information infrastructure of modern social movements," *Journal of Communication*, vol. 62, no. 2, pp. 315–329, 2012.

[200] D. Gayo-Avello, "A meta-analysis of state-of-the-art electoral prediction from twitter data," *Social Science Computer Review*, p. 0894439313493979, 2013.

[201] F. D. Carpentier, S. Knobloch, and D. Zillmann, "Rock, rap, and rebellion: Comparisons of traits predicting selective exposure to defiant music," *Personality and individual differences*, vol. 35, no. 7, pp. 1643–1655, 2003.

[202] P. Fischer, E. Jonas, D. Frey, and S. Schulz-Hardt, "Selective exposure to information: The impact of information limits," *European Journal of Social Psychology*, vol. 35, no. 4, pp. 469–492, 2005.

[203] M. Chen and J. A. Bargh, "Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus," *Personality and social psychology bulletin*, vol. 25, no. 2, pp. 215–224, 1999.

Appendix A

PRE-SELECTED AND CHOSEN TOPICS IN STUDIES 1-3 (CHAPTER 3-5)

Table A.1: **Controversial topics used as experimental tasks in Studies 1-3**

Controversial topic	Chosen in Study 1 (Chapter 3)	Chosen in Studies 2 & 3 (Chapters 4 & 5)
Should certain vaccines be required for children?		
Should euthanasia or physician assisted suicide be legal?	✓	✓
Is drinking milk healthy for human?		
Should prescription drugs be advertised directly to consumers?	✓	✓
Do violent video games contribute to youth violence?	✓	✓
Should certain performance enhancing drugs (such as steroids) be accepted in sports?	✓	
Should all Americans have the right be entitled to health care?		
Should people become vegetarian?	✓	✓
Should death penalty be allowed?	✓	✓
Should the US have sent troop to Iraq?	✓	
Are the 2010 health care reform laws good for America?		
Is human activity a substantial cause of global climate change?		
Should Social Security be privatized?	✓	✓

Appendix B

EXAMPLE OF THE TOPIC QUESTIONNAIRE IN STUDY 1 (CHAPTER 3)

1. Legalizing euthanasia is :

Harmful	1—2—3—4—5—6—7	Beneficial
Unethical	1—2—3—4—5—6—7	Ethical
Unnecessary	1—2—3—4—5—6—7	Necessary
Wrong	1—2—3—4—5—6—7	Right
Undesirable	1—2—3—4—5—6—7	Desirable

2. How certain are you about your attitude on this issue?   3. How

Uncertain   1—2—3—4—5—6—7   Certain

important is your point of view on this issue to you?

Little   1—2—3—4—5—6—7   A lot

4. How much is the topic related to your core value?

Little   1—2—3—4—5—6—7   A lot

5. How much do you know about this topic?

Little   1—2—3—4—5—6—7   A lot

6. How interested are you in knowing more about this topic?

Little   1—2—3—4—5—6—7   A lot

7. How much do you desire to know the truth about the topic regardless of your own point of view?

Little   1—2—3—4—5—6—7   A lot