Unbiased Learning of Controversial Topics

V.G.Vinod Vydiswaran[1], ChengXiang Zhai[2], Dan Roth[3], Peter Pirolli[4]
[123]University of Illinois, 201 N.Goodwin Ave, Urbana, IL 61801, USA
[4]Palo Alto Research Center, 3333 Coyote Hill Road, Palo Alto, CA 94304, USA
[1]vgvinodv@illinois.edu, [2]czhai@illinois.edu, [3]danr@illinois.edu, [4]Peter.Pirolli@parc.com

ABSTRACT
When presented with many relevant documents about a controversial topic, humans do not always read and trust them uniformly. Instead, they tend to follow and agree with articles and sources that hold similar viewpoints as theirs, a phenomenon known as confirmation bias. This suggests that when acquiring additional knowledge about a controversial topic, human biases and viewpoints about the topic may affect what information is considered trustworthy. We designed a user study to analyze various factors that may help in learning about a controversial topic without bias. In this paper, we study the impact of presenting contrasting viewpoints on learning about controversial claims. Our analysis shows that exposing subjects to contrasting viewpoints helped them learn more efficiently and reduce strongly-held biases.

Keywords
Claim verification, unbiased learning, user study

INTRODUCTION
The Internet has become one of the primary sources of information in a variety of domains. Online news portals are gaining popularity steadily over the last decade, and traditional print media is losing ground. Patients and caregivers search online for information about particular diseases, share their medical history, and learn about treatment options through web portals and health forums. Students rely on online resources to complete assignments in history, literature, and other subjects. In all these tasks, naïve information seekers assume that the information available online is accurate, trustworthy, and unbiased. However, the Web is a mix of both well-curated, edited content and unchecked, freelance content. Although the former is considered more trustworthy than the latter, it is possible that some reputable sources have biased opinions and support false claims. If online information seekers relied only on these sources, they may get a biased view.

Consider a scenario where Alice, an Internet surfer, wants to know how healthy the food in her child's school cafeteria is, and specifically if chocolate and flavored milk provided in schools is a healthy food choice for her child. Depending on the keywords she chooses to search about the topic, she might encounter news about a recent ban of chocolate milk in certain schools, or learn about health benefits of milk in growing kids. She might find results from news media organizations reporting on the ban, news reports on activist groups actively encouraging milk consumption, or even questions and answers from concerned parents interacting via community-driven question answering services.

We wanted to study various factors that enable humans to acquire additional information about controversial topics in an unbiased fashion. We designed BiasTrust, a user study to understand if the display of contrasting viewpoints and source expertise helped users in this task. In this paper, we present results gathered from the pre-study and post-study surveys and feedback interviews, and show that subjects learnt more about the controversial topics and reduced their strong biases when contrasting viewpoints were shown.

RELATED WORK
Understanding what documents people read is related to research in many fields. Psychologists have studied the phenomenon of confirmation bias (Plous, 1993; Baron, 2000), which states that people tend to favor information that confirms their beliefs, both in choosing what to read, and in accepting or rejecting what they read. Similarly, researchers in political science (such as Taber and Lodge, 2006) studied how people process information in a biased fashion, i.e. how they uncritically accept supporting claims, but argue against claims contrary to their beliefs.

Researchers have looked at aggregating information from multiple sources to answer specific questions. Wu and Marian (2007) studied how to collect information from multiple sources over the Web, specifically to find answers using corroborative evidence from multiple sources.

ASIST 2012, October 26-30, 2012, Baltimore, MD, USA.

Gallard, Abiteboul, Marian, and Senellart (2010) looked at collecting information from contrasting viewpoints. We consider the next step, of how to present contrasting viewpoints to enable users learn about the topic efficiently.

Researchers have also studied factors that influenced what information users accessed and how they processed them. This includes work on building tools to increase transparency and credibility of Wikipedia articles (Pirolli, Wollny, & Suh, 2009; Su, Chi, Kittur, & Pendleton, 2008) that give users a clearer sense about what information is credible and what is not. Pariser (2011) investigated the notion of *filter bubble*, in which search engine personalize results in such a way that users get to see only articles that are similar to what they have already seen. This not only encourages confirmation bias, but also hides contradictory viewpoints. Our study tries to understand how to overcome these shortcomings by presenting contrasting, yet credible evidence to users. We believe that our work is the first to study how contrasting viewpoints and source expertise rating help users learn about controversial topics.

BIASTRUST: STUDY DESIGN

Deciding on truthfulness of controversial claims is basically a learning task, where an inquisitive user tries to learn as much as possible about the claim and assimilate all evidence in support of or against the claim. However, as we pointed out earlier, previous research by psychologists and others have shown that users tend to access information that supports their own viewpoints. We wanted to study if presenting contrasting viewpoints side-by-side helps users overcome this bias and learn more effectively.

We designed a system that retrieves relevant, trustworthy documents about a controversial topic and helps users get an unbiased perspective about the topic. In order to optimize the design of how to present credible information to users, we conducted a user study called BiasTrust. The user study investigates various factors that may affect which documents humans read and which ones they judge as relevant and credible. This study would help us design and improve interfaces for an automated claim verification system that allows users to validate claims by reviewing evidence for and against the claim.

The user study was designed as a learning task, where users were asked to learn as much as possible about a topic within a stipulated time. This setup helped users decide what is important for them to read, given the limited time, and choose documents accordingly. We could then observe their actions and study how various factors helped or hindered them in this learning process. By choosing controversial topics where there is a genuine evidence both supporting and opposing the claims, we could understand how preference-based factors affect the learning process.

The study was conducted in four stages, viz. (i) Pre-study survey questionnaire, (ii) Study phase, (iii) Post-study questionnaire, and (iv) Feedback interview. The first three stages were conducted online, while the feedback interview was face-to-face. These four stages are explained below.

Stage 1: Pre-study questionnaire

The pre-study questionnaire was designed to measure the subjects' prior knowledge about the controversial topic. Specifically, subjects were asked a few questions to help us gauge their (lack of) knowledge and bias towards/against important issues relevant to the topic.

Subjects answered the questions on a four-point Likert scale. For the knowledge-related questions, they were asked to indicate how much importance they gave to a few sub-topics relevant to the topic of interest. The answers on the Likert scale ranged from (i) 'insignificant' to (iv) 'very significant'. For bias-related questions, subjects were asked about their opinion/preference for particular concepts. The answers could range from (i) 'strongly against' to (iv) 'strongly in favor of' the issue. For all questions, the subjects could also select the 'I don't know' answer if they did not know enough about the sub-topic being discussed.

For example, if the issue being discussed is whether drinking milk is healthy for humans, the questionnaire would include questions asking subjects if they were aware of the issue of flavored milk being distributed in schools (knowledge question) and if they believed flavored milk is a healthy choice (bias question). By using knowledge and bias related questions and limiting the nature of allowed responses, we encouraged subjects to think about their positions on many sub-topics related to the overall issue.

Stage 2: Study phase

Once the subjects' responses to the pre-study questionnaire were recorded, they were directed to the study phase. For each relevant passage, subjects were shown the source of the passage, the sub-topic the passage is closely related to, and whether the passage was in favor or against the sub-topic. In some interface variants, the contrasting viewpoints were shown side-by-side; while in others, users had to pro-actively opt to see a passage from the contrasting viewpoint as the next result. In some variants, the expertise rating of the source was also shown to further help subjects decide if they wanted to read the passage.

For each passage that the subjects read, they were asked to answer three questions about it, viz. did they agree with what was being said in the passage; did they get any new information; and did they believe the information was biased with respect to the topic being discussed. These three questions allowed us to quantify the perceived importance of the passage. By first observing the subjects choose which passages to read, and then analyzing their opinion about the quality of the passage based on agreement, novelty, and perceived bias, we could record how the passages helped them learn about the topic.

Subjects were asked to continue reading in this fashion until they believed they had read enough about the topic.

Stage 3: Post-study questionnaire

Once the subjects decided to quit the study phase, they were asked to respond to a series of questions about the topic. This was similar to the set of topic-specific questions that were posed during the pre-study questionnaire. The subjects then wrote a short summary essay on what they learned. They were also asked to give feedback on what interface features helped them in their task and to suggest additional features they would like to see in such a system.

Stage 4: Feedback interview

The final stage of the study was a face-to-face meeting and debriefing session, conducted within a week of completion of the online phases. Participants were debriefed about the study and were informed about the factors that were being studied. Participants could also provide additional feedback about the system and suggest changes to improve the study.

SETTING UP THE USER STUDY

Data and Study topics

We enabled the study for two controversial issues, one from the health domain and the other from politics. The primary claims ("issues at hand") corresponding to the topics were:

- *Milk* issue: Drinking milk is a healthy choice for humans.
- *Energy* issue: Alternate sources of energy are viable alternatives to fossil fuels.

We chose these particular issues because they are fairly similar in terms of biases they may invoke. We wanted topics that are controversial, but also have scientific evidence to justify either viewpoint. Notably, these issues were different from other controversial issues, such as euthanasia and right to life, that may provoke strong emotional biases that are hard to overcome.

For each of the two issues we included in the study, we collected over 350 snippets of text from ProCon.org[1], a nonpartisan, non-profit public charity website. The website collects quotes and passages from people, organizations, websites, and other sources relevant to the issue being discussed. It groups the quotes based on relevant questions or sub-topics within the issue, and categorizes them as *pro* (in favor of the question being asked), *con* (against the question being asked), or neither *pro* nor *con*.

User Profile and Interaction Summary

We invited volunteers to participate in the BiasTrust study by announcing the study on mailing lists in many departments within a large, diverse, public university; and to members of the larger local community not affiliated to the university. In all, 24 volunteers participated in the study, and the average age of participants was 28.6 ± 4.9 years. Overall, we collected information from 40 study sessions, since most participants chose to take part in both tasks.

[1] ProCon.org, http://www.procon.org/

The profile of how subjects interacted with the system was similar for the two tasks. Typically, subjects took 7–10 minutes to complete each of the pre-study and post-study questionnaires. On-an-average, they spent 26.5 minutes in the study phase for each topic. During this time, they read an average of 18.6 documents per session, but considered as many as 31 documents, including those they chose to skip.

OBSERVATIONS AND RESULTS

Extent of Learning

In this paper, we look at the learning task, and analyze if exposure to contrasting viewpoints helped or hindered learning about the controversial topics. We were able to capture this based on subjects' responses to the same questions in the pre-study and post-study questionnaires.

Subjects learned about topics they did not know

In our study, the subjects tended to read more on topics they did not know about, rather than read about topics they already knew. In all, subjects self-reported not knowing about a particular phenomenon or sub-topic on 86 occasions (in about 10.8% cases). Out of these, the same subjects later reported to have learned something about the sub-topic in 63 instances. This constitutes a learning rate of 73.26%.

We also observed that subjects were inquisitive about the contrasting viewpoints, and read them both when the passages were shown right next to each other. In Vydiswaran, Zhai, Roth, and Pirolli (2012), we present additional results on how explicitly showing contrasting viewpoints and source expertise rating increases readership. We conclude that these helped subjects overcome the presentation bias.

Subjects changed strongly-held biases

The study also helped the subjects in moderating strong biases towards the issues. At the start of the *Milk* study task, subjects were asked if they considered milk to be a healthy choice for humans. Most subjects overwhelmingly believed it to be a very healthy choice. The average rating for milk being a healthy food choice in pre-study survey was 4.55 ± 0.59 on a [1–5] Likert scale. However, after being exposed to evidence about possible contamination of milk, added chemicals, and adverse impact on health for certain individuals, the average rating for milk as a healthy food choice in the post-study survey reduced to 3.91±1.08. This is a statistically significant change in previously held belief. We also observed that inquisitive users who read more contrasting viewpoints had a higher tendency to change their biased opinion. Many subjects also mentioned this explicitly in the essays they were asked to write in the post-study stage. One subject wrote: "*I did not know that milk had so many worrisome factors caused due to mass-scale production. I have to be more careful!*"

On the *Energy* task, subjects were asked if they believed alternate energy sources were viable alternatives to fossil fuels. On-an-average, we did not find an overwhelming bias

for this issue. The average rating for this question in the pre-study survey was 2.80 ± 0.51 on a [1-4] Likert scale, which means most people believed that alternate energy sources could replace just about significant portion of power generated by fossil fuels. In the post-study survey, we find that the optimism increased, but only slightly. The average rating in post-study survey was 2.98 ± 0.55.

Issue at hand	Measure	I	D	Change
Type of questions: Bias questions				
Milk	Variance	2	9	-31.0%
Energy	Variance	2	5	-27.9%
Type of questions: Knowledge questions				
Milk	Mean	7	2	+12.3%
Energy	Mean	8	5	+3.3%

Table 1. Effect of the study phase on responses to knowledge and bias questions. I (D) denotes number of questions in which the measure increased (decreased).

Learning about sub-topics relevant to the study task

Next, we studied how reading about the sub-topics helped the subjects learn more about the task. For each task, we had twenty questions that were either knowledge-oriented or bias-oriented about specific sub-topics. We found that subjects changed their opinion on various sub-topics on 285 instances. Out of these, in 40.8% cases, they changed the importance they gave to the sub-topics significantly (by at least one point on a [1-4] Likert scale). Similarly, in 24.3% cases, they reported to have reduced the bias strength by at least one point, moving away from extreme bias positions.

We wanted to quantify the increase in the average knowledge rating (mean) and the reduction in the spread of the bias rating (variance). Table 1 summarizes the relative shift in responses to knowledge and bias questions.

In the *Milk* task, there were nine knowledge related questions, seven of which got an overall higher rating and the remaining two got a poorer rating. In all, the average knowledge rating increased by 12.3%. Further, out of the eleven bias related questions, the variance reduced in nine and increased in the other two, with an average reduction in variance of 31.0%. Both changes are statistically significant at p=0.05 level using Wilcoxon's signed rank test.

Similarly, for the *Energy* task, we found that the variance reduced in five out of the seven bias-related questions, with an average reduction of 27.9%. Finally, although the knowledge rating increased in eight out of thirteen knowledge questions, the average increase of 3.3% was not found to be statistically significant. On further analysis of the responses, we found that many subjects read only about a few sub-topics, i.e. they did not read about most alternate energy sources. So, because of limited exposure, their opinion about viability of alternate energy source replacing fossil fuels did not change significantly. However, subjects with a strong bias against the issue demonstrated increased knowledge about the viability of alternate energy sources, and reduced their bias after the study phase.

CONCLUSION AND FUTURE WORK

In this paper, we studied the factors that help users learn about controversial topics via a user study. We find that showing contrasting viewpoints helped them significantly reduce strong biases in favor of or against topics and helped them learn about new sub-topics in an unbiased fashion. Going forward, we believe that the insights gathered from this study will help us build more effective interfaces that help users learn about and validate controversial claims.

ACKNOWLEDGMENTS

This research is partially funded by grants from the MIAS Center of Excellence of the Department of Homeland Security, the Information Trust Institute, the Army Research Laboratory, and the Office of Naval Research.

REFERENCES

Baron, J. (2000). *Thinking and deciding*. Cambridge Press.

Gallard, A., Abiteboul, S., Marian, A., & Senellart, P. (2010). Corroborating information from disagreeing views. In *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining (WSDM)* (pp. 131-140). ACM.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin Press.

Pirolli, P., Wollny, E., & Suh, B. (2009). So you know you're getting the best possible information: A tool that increases Wikipedia credibility. In *Proceedings of the 27th annual SIGCHI Conference on Human factors in Computing Systems* (pp. 1505-1508). ACM.

Plous, S. (1993). *The psychology of judgment and decision making*. McGraw-Hill.

Suh, B., Chi, E. H., Kittur, A., & Pendleton, B. A. (2008). Lifting the veil: Improving accountability and social transparency in Wikipedia with WikiDashboard. In *Proceeding of the 26th annual SIGCHI Conference on Human factors in Computing Systems* (pp. 1037-1040). ACM.

Taber, C. S. & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science 50(3)*, 755-769.

Vydiswaran, V., Zhai, C., Roth, D., Pirolli, P. (2012). BiasTrust: Teaching biased users about controversial topics. In *Proceedings of the 21st International Conference on Information and Knowledge Management (CIKM)*. (Under print) ACM.

Wu, M. & Marian, A. (2007). Corroborating answers from multiple Web sources. In *Proceedings of the 10th International Workshop on Web and Databases (WebDB).* (pp. 1-6).