A Crowdsourced Study of Visual Strategies for Mitigating Confirmation Bias

Tee Chuanromanee
Computer Science and Engineering
University of Notre Dame
Notre Dame, USA
tchuanro@nd.edu

Ronald Metoyer
Computer Science and Engineering
University of Notre Dame
Notre Dame, USA
rmetoyer@nd.edu

***Abstract*—Confirmation bias is a type of cognitive bias that involves seeking and prioritizing information that conforms to a pre-existing view or hypothesis that can negatively affect the decision-making process. We investigate the manifestation and mitigation of confirmation bias with an emphasis on the use of visualization. In a series of Amazon Mechanical Turk studies, participants selected evidence that supported or refuted a given hypothesis. We demonstrated the presence of confirmation bias and investigated the use of five simple visual representations, using color, positional, and length encodings for mitigating this bias. We found that at worst, visualization had no effect in the amount of confirmation bias present, and at best, it was successful in mitigating the bias. We discuss these results in light of factors that can complicate visual debiasing in non-experts.**

***Index Terms*—visualization, crowdsourcing, confirmation bias**

I. Introduction

Cognitive biases are errors in thinking that affect judgment and decision-making because humans tend to use heuristics or "mental shortcuts." These shortcuts allow for faster and more efficient information processing [5]. However, utilizing heuristics in higher-stakes environments, where more attention to detail and careful consideration are needed, can lead to partiality and bias [6].

The dual systems theory [26] explains why cognitive bias manifests. This theory frames the decision-making process into two modes of thinking. The first mode, System 1, is automatic, intuitive, and fast, while System 2 is conscious, reflective, and slow. System 1 is associated with using heuristics, shortcuts and cognitive biases. Thus, in order to mitigate biases, engaging System 2 is helpful because it elicits a slower, more conscious thinking process [9].

Confirmation bias is a type of cognitive bias that leads to prioritizing or seeking evidence that affirms existing hypotheses, views, or expectations. It is present in many domains [22] including healthcare [1], [31], suspect interviews [14], political attitudes [17], and business [23], and can lead to discrimination [13] and misdiagnosis [21]. Thus, it is important to be able to detect and mitigate confirmation biases in domains involving decision-making and sensemaking. Several conditions can lead to an increased occurrence of confirmation bias in such processes. For example, tasks involving time constraints [29] and large amounts of information [27] involve significant cognitive strain [29] leading to ideal conditions for confirmation bias to manifest.

Here, we explore the use of visual representations to draw attention to evidence polarity (i.e., positive, negative, neutral) to engage System 2 thinking in decision-making domains. While visualization methods may be appropriate for mitigating some biases such as availability bias [28], it is not clear how they might affect other types of cognitive biases such as confirmation bias. Thus, our research questions are as follows:

RQ1: How do visualizations in general affect the phenomenon of confirmation bias?

RQ2: What visual representations are more effective at bias mitigation?

To explore these questions, we investigated the use of visualization for mitigating confirmation bias through an Amazon Mechanical Turk (MTurk) experiment. We found that visualization either makes no difference in the level of confirmation bias or may actually mitigate this bias.

We contribute an exploration of the role that visualization plays in the manifestation and mitigation of confirmation bias and an examination of factors that can complicate visual debiasing in non-experts. In the following sections, we present literature on confirmation bias and bias mitigation before walking through our study. We then discuss our results and their implications for the use of visual cues with non-experts.

II. Background and Related Work

Many factors can affect the amount of confirmation bias present, making this a complex phenomenon to study. There are two main schools of confirmation bias mitigation techniques. The first aims to directly mitigate the bias using education and training. However, these techniques create additional cognitive load [18]. The second approach aims to change the environment "to fit the decision-maker's cognitive processes" [2], [16]. This is more suitable for applications where cognitive strain can pose a problem. Successful examples include showing alternate points of view [19], using games [10], and giving participants enough time to search for information [25].

This material is based upon work supported by the National Science Foundation under Grant No. 1816620.

978-1-6654-4214-5/22/$31.00 ©2022 IEEE

TABLE I
SELECTED EVIDENCE AND RATINGS FOR POSITIVE HYPOTHESIS IN THE PUBLISHING DOMAIN. RATINGS WERE GENERATED BY THE RESEARCH TEAM.

Evidence	**Supportiveness** (7:most,1:least)
Critics fault the writer's previous works for being dry and boring.	1.75
The proposal lacks detail about the plot of the book.	2.65
The genre of the proposed book (Sci-fi dystopia) is slightly different from the ones Calle Publishing is known for (Sci-fi fantasy).	2.75
The writer's agent has negotiated a higher rate including royalties that could cost the publisher money if the book does not sell well.	3.10
The book proposal has potential for it to be a series.	4.45
The writer has a devoted following on social media.	4.75
The concept is new and original, which is something Calle Publishing is looking for.	6.25
The proposal promises that the book will be deep and thought-provoking.	6.40

These debiasing methods tend to be quite general and not necessarily appropriate for specific domains. Thus, other studies investigating confirmation biases tend to come from domains such as intelligence analysis [8], [30].

In particular, Cook and Smallman evaluated confirmation bias in the context of intelligence analysis [8]. They presented participants with a hypothesis and eight pieces of evidence, with each piece of evidence containing a rating between one (least supportive of hypothesis) and seven (most supportive of hypothesis). They asked participants to select and rank the helpfulness of four pieces of evidence that were the most useful in helping them evaluate the hypothesis. In this study, confirmation bias is manifested when (1) participants select more confirming evidence (described as those rated above four in the Likert scale) as useful, and (2) of the selected evidence, more confirming evidence was ranked highly than compared to less-confirming evidence [8].

They found that using visualization with a spatial encoding of the evidence distribution placed on a horizontal axis mitigated confirmation bias compared with a control group that was presented with text-only information. However, both groups still displayed confirmation bias so the reduction of bias was not complete [8]. This experiment provides an appropriate setting for a holistic review environment, therefore, we sought to apply their domain-specific experiment to a book publishing domain in order to investigate the debiasing potential of visualizations.

III. STUDY AND RESULTS

Our study is based on that of Cook and Smallman [8], but with minor modifications to better fit our domain of interest. In particular, to determine that confirmation bias exists and whether using visualization will mitigate confirmation bias, we designed a Mechanical Turk study testing confirmation bias based on format of the evidence (with six levels, including text-only vs. each of five different visualizations used) and type of hypothesis (positive vs. negative).

A. Design

The factors in our 2x6 experiment were the Hypothesis used and the types of Visualizations displayed. The hypothesis presented has two levels, which can be Positive (e.g. "This is a strong book proposal"), or Negative (e.g. "This is a weak book proposal"). The format of the information presented can be Text-only or include one of five Visualizations. In each condition, we had 45 participants for a total of 540 participants. We required that participants had an HIT approval rate greater than or equal to 95%, had completed more than 5,000 HITs, and were college graduates.

B. Stimuli and Procedure

Our participants were presented with stimuli in the form of fictitious scenarios (similar to [8]) intended to be representative of real-world situations with regard to topic and decision-making strategy. The scenario required a cognitively complex decision-making task that involved selecting evidence in support of or against a decision hypothesis.

After obtaining informed consent, we presented an introduction to the scenario and the context of the domain that would help participants gain background information in order to make informed decisions in the tasks. We then described the task to participants – that they needed to decide whether or not to accept or reject the decision hypothesis.

To help participants calibrate themselves to the goals and values that were relevant to the domain, we included two "profiles" of sample information – one that supported the hypothesis and one that contradicted the hypothesis. After reviewing the instructions, participants would begin the experiment where they were presented with the hypothesis and the evidence to select from and rank. We told participants to assume the role of a decision-maker (i.e., an editor), and to also justify their choices through a text box. Note that we asked participants to determine whether the hypothesis was true or not, in addition to determining which pieces of evidence were supportive of their decision. This was an attempt to encourage participants to examine contradicting evidence and not simply to look for supportive evidence. We gave participants an hour to finish the HIT, however, most participants required approximately 20 minutes to complete the task. This experimental setup is provided in the supplemental materials.

We hypothesized that confirmation bias would be observed in all groups. We also expected that visualization would reduce the confirmation bias in all cases given the prior results of Cook and Smallman [8]. However, we did not expect that this difference would be pronounced when comparing types of hypotheses (that is, there would be no interaction between Visualization and Hypothesis factors).

C. Presented Evidence

Through a series of pilot studies, we determined the best way to both present the scenario to the participants and to represent

the visualizations. The studies are performed in a complex decision-making domain, a book publishing scenario. In this scenario, Calle Publishing is a highly regarded fictional book publishing company. The participant is tasked with analyzing pieces of evidence from a book proposal, given a hypothesis about the proposal (i.e., whether the proposal was strong or weak) and ratings for each piece of evidence.

We calibrated the evidence supportiveness ratings in order to provide a level of standardization for non-experts. In both the positive and negative hypothesis conditions, evidence that supports the given hypothesis correspond to higher ratings than those that tend to refute the hypothesis.

To create the evidence provided to the participants in this publishing scenario, the team rated the evidence individually, then calculated the interclass correlation to be 0.844. We then computed the average of the assigned evidence. The evidence and associated ratings are presented in Table I.

D. Visual Representations

We utilized five different visual representations in the study. The first visualization contains a color encoding to highlight the evidence ratings as seen in the "Evidence" section in Figure 1. Each color corresponded to a point on the given Likert scale of supportiveness of the evidence. We hypothesized that a compact color encoding would help participants be aware of the distribution of evidence.

Because position encodings are known to be more effective at conveying quantitative values than color encodings [20], we implemented a simple position slider for our next visualization which shows whether the evidence they selected and ranked seemed to support or refute the hypothesis, and changes whenever participants selected or ranked different pieces of evidence. Figure 2 shows the position slider when the weighted value of the selected evidence leans towards refuting the hypothesis. The position of the dot is computed using a weighted average of the selected evidence ratings.

In the third visualization (as seen in Figure 3), we highlighted only the refuting information using a single color and removed the legend. We did this specifically to challenge participants' mental models (such as in Choi et al.'s work [7]), draw their attention to the refuting evidence, and encourage them to consider these pieces of evidence further.

Next, we tested the effectiveness of a length encoding of the ratings using bar sparklines (as seen in Figure 4) as they are compact and would be easy to implement in real-world situations where space can be limited. Length encodings should also be effective to communicate and compare the ratings [20].

Finally, we tested the radar chart. We explored it because it is a useful way to *compare* multiple sets of data which would be a reasonable use case in a real-world application setting. The radar chart visualization indicated the neutral rating of 4.0 by using bolded font. The the eight pieces of evidence we presented to the participants were labeled using the letters A to H. We also told participants that the larger the filled area was, the more supportive the evidence shown will be.

TABLE II
AGGREGATE RESULTS FOR MAIN EFFECTS FOUND IN THE MTURK STUDY (BOLDED VALUES ARE SIGNIFICANT). NOTE THAT THIS STUDY USED THE BOOK PUBLISHING DOMAIN.

Condition	Mean Rating			Comparison to Text-Only	
	Negative	**Positive**	**Both**	**F(1,178)**	**p**
Text-only	4.49	5.30	4.90	-	-
Color visualization with legend	4.51	5.03	4.77	0.863	0.39
Highlight refuting evidence only	**4.14**	**4.93**	**4.54**	**2.48**	**0.01**
Position slider	4.40	4.92	4.66	1.60	0.11
Sparklines	4.76	5.31	5.04	-0.95	0.35
Radar Chart	4.44	5.03	4.74	1.08	0.28

E. Results

To compare the effectiveness of the five visualizations, we performed a study using the publisher's domain comparing the text-only case with the visualizations described above. To analyze the data, we utilized a repeated-measures ANOVA, and report the weighted means of the selected evidence rating, p-values, and F-values. Figure 5 and Table II show the results.

Bias was present in all cases. Utilizing an ANOVA of the weighted rating means, we found a significant main effect of both Visualization ($F(5, 534) = 2.87$, $p = 0.014$) and Hypothesis ($F(1, 538) = 55.22$, $p < 0.001$). However, there was no Visualization x Hypothesis interaction ($F(5, 534) = 0.43$, $p = 0.838$). In examining our significant Hypothesis main effect, we found that the Positive hypothesis had a greater mean rating of 5.09 compared with the Negative hypothesis, which had a mean rating of 4.46.

Because we found a Visualization main effect, we ran some contrasts to compare the difference in the weighted means between the text-only condition and each of the visualizations. We found that only the comparison between the text-only condition and the highlighted refuting evidence were significantly different, as seen in Table II.

IV. DISCUSSION

Our results raise many questions on the role of visualization and different debiasing techniques in confirmation bias mitigation. We found that visualizations were just as effective as text representations (no worse in terms of confirmation bias) and may potentially help mitigate confirmation bias. In the following sections, we discuss these results and provide potential explanations and how these results may apply beyond MTurk workers to domain experts.

A. General and Visualization Debiasing Strategies

Our study showed that general debiasing strategies that are implemented using visualizations can be effective. Combining general and visualization debiasing strategies may improve the level of debiasing, however, this needs to be validated.

There is a difference between using a general debiasing strategy and using a visualization debiasing strategy. A general

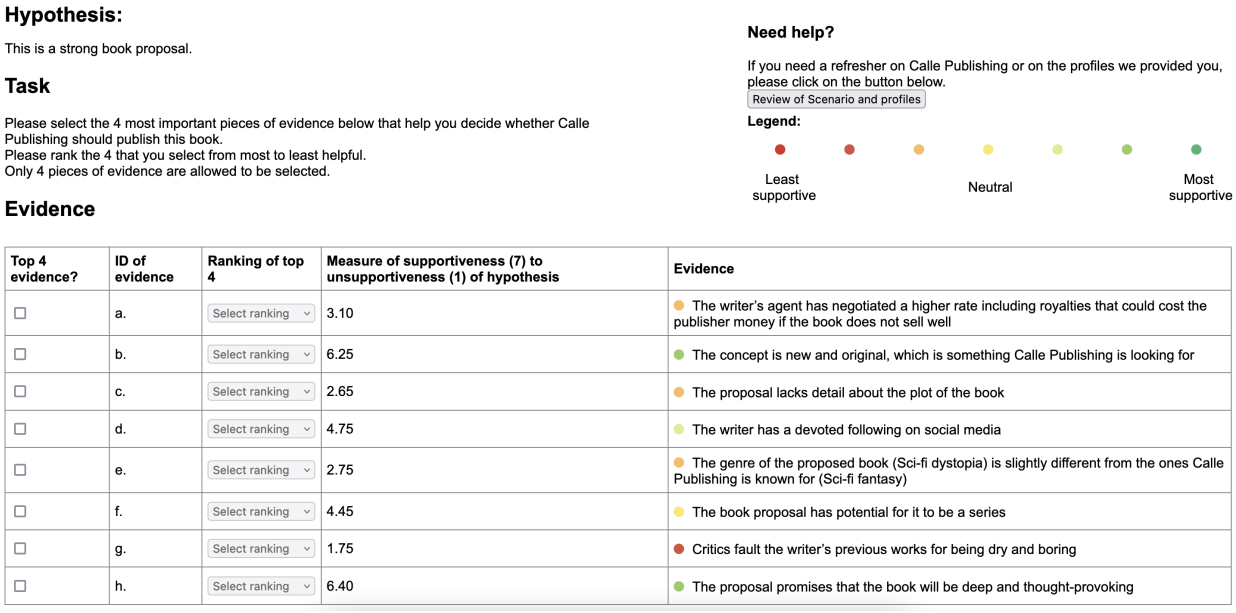

Fig. 1. MTurk study with color visualization and legend as seen in the positive hypothesis condition. This figure also shows the way we presented the evidence to the participant. All conditions will contain the same basic structure, but with different visualizations.

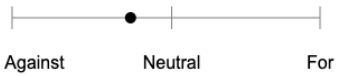

Fig. 2. The simple slider visualization shown here is placed above the evidence that is presented to the participants. The value of the slider shows a lean against the given hypothesis.

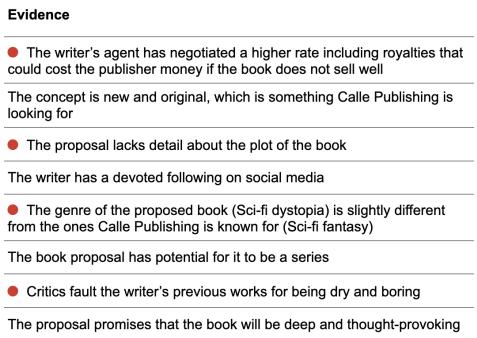

Fig. 3. The visualization where we only highlighted the refuting evidence. The red dots here highlight the information that refutes the given hypothesis (i.e. that this was a strong book proposal).

debiasing strategy is simply that which comes from the non-visualization bias literature. It can be applied both in text and visualization forms. If visualization is used, it is used as an application of that general strategy. An example would be highlighting specific information in order to bring attention to the alternative hypothesis.

In contrast, using a visualization-based debiasing strategy means that the use of visualization itself is meant to debias [24]. An example is the use of a dashboard to present data otherwise distributed across the interface in a compact

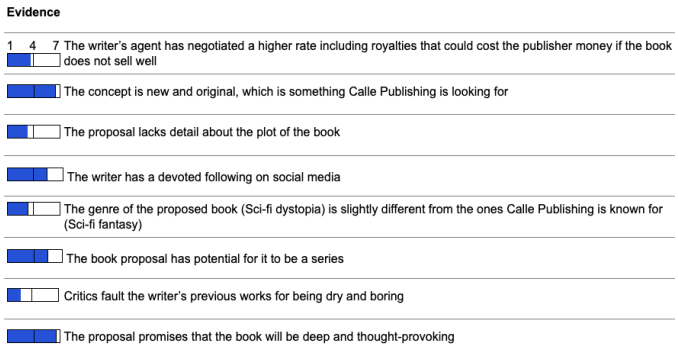

Fig. 4. The sparklines visualization used in the study. Each sparkline has a vertical line in the middle to indicate the neutral position. The larger and longer the blue filled area is, the more supportive the evidence. The figure shoes the positive hypothesis.

form, thus combating availability bias, where easily recalled information is over-weighted in the decision-making process.

It is possible to combine the two strategies. For example, suppose we aim to mitigate the presence of availability bias in the college admissions process [29]. Availability bias manifests if reviewers consider and put more weight on easily-remembered or easily-accessed pieces of information. This could be mitigated by a combination of the two strategies. A general strategy would be to use a color encoding to draw attention to potentially missed information, while a visualization strategy would be to encode the application data itself in such a way that it is more easily available (e.g., showing all test scores in a chart).

In our case, we only utilized general strategies even though we used visualization to implement those strategies. For

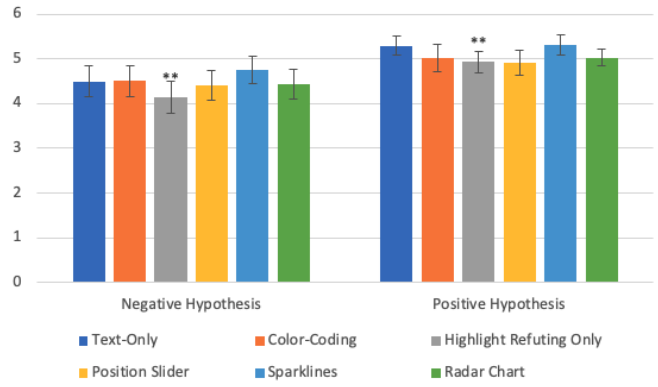

Fig. 5. Mean ratings for each condition in the study. Error bars represent a 95% confidence interval. The condition where we highlighted refuting evidence only has a significant difference in rating compared to the text condition ($p < 0.05$). While we separate our results by the hypothesis used, we did not encounter a Visualization x Hypothesis interaction effect.

example, when we highlighted the refuting information only, we used visualization as a tool to draw attention to pieces of information that may have been missed. The slider is also a general strategy where visualization presents additional information which shows the decision-maker their leaning and encourages introspection, switching from System 1 to System 2 cognitive processes [29].

Our results suggest that visualizations will perform similarly to text representations and at best, may even serve to mitigate confirmation bias (**RQ1**). This means that for trained reviewers, visualization is likely safe to use and perhaps advantageous. It has the potential to not only mitigate confirmation bias but also ease cognitive load [4] and combat other biases such as availability bias [28]. However, more empirical evidence is needed in terms of *what* makes a debiasing method, both visualization-based and general, effective.

In our studies, highlighting only refuting evidence showed the most promise at being able to mitigate bias. However, we also saw that the sparklines potentially increased the level of confirmation bias. Amplification may occur due to the presence of other forms of biases.

B. Beyond Confirmation Bias

Cognitive dissonance is another potential factor that could help explain our results. It is known that cognitive dissonance increases confirmation bias [12], [15]. We suspect, however, that visualizations can help participants resolve their dissonance in a way that tends to lean towards debiasing.

Cognitive dissonance refers to the conflict experienced when opposing pieces of evidence cause people to feel uneasy, and tends to lead to confirmation bias. The uneasiness is usually resolved by weighting one piece of evidence more favorably than the other [3] because it is the easier way to resolve the dissonance [11]. Our studies manufactured cognitive dissonance among participants because while we presented the participants with balanced evidence (that should indicate a mediocre book proposal), we told participants to evaluate a hypothesis (which essentially told them to lean one way or another). Another manifestation of cognitive dissonance could be between participants' opinion of the evidence and the hypothesis that we presented to them.

Cognitive dissonance is a state of mind that people tend to avoid, and the easiest way to reduce the feeling of dissonance is by choosing only the evidence that is consistent with the leaning of the decision about to be made as well as ignoring the evidence that is inconsistent [11]. Quantitatively, cognitive dissonance leads to a greater distance from a rating of 4.0 (whether this distance is to the positive or negative direction).

Cognitive dissonance is made worse by increased cognitive load [11]. Visualization, we suspect, decreases this load and makes data more salient helping them resolve dissonance faster by drawing a person's attention more quickly to the information that can help them resolve their dissonance. This is one explanation for why highlighting evidence was the most successful, since it does not introduce much cognitive load, compared to other visualizations such as the radar chart. We suspect this effect would be less pronounced with professional reviewers trained to mind their personal (and cognitive) biases; this group will be considered in future work.

C. Limitations

Isolate various forms of potential bias, such as personal bias, is difficult, thus we understand that it is impossible to measure this all potential biases. Our study design also relies on a presumed hypothesis that is provided to the participant. This was necessary since it is challenging to create a situation in which the participant develops their own hypothesis. We understand that in this case, the participants views may not be anchored in the presumed hypothesis, thus presenting a threat to ecological validity. In addition, because our studies focused on non-experts, we cannot generalize our findings to experts.

Furthermore, confirmation bias can manifest in many forms. For example, congruence bias causes participants to only test if a hypothesis is true or not and not test any alternative hypotheses [32]. This manifested in our studies, for example, when participants could have solely considered whether a book proposal is strong, but not whether the proposal is a weak or even a mediocre one.

V. Conclusion

Through a series of MTurk studies, we found that visualization can be effective in mitigating confirmation bias. However, for non-experts, factors such as cognitive dissonance may complicate debiasing. Thus, a multidimensional approach is needed to reduce confirmation bias in non-experts.

Because using visualizations provides multiple advantages apart from debiasing, designers should carefully use visualizations when appropriate. The visualizations we used were manifestations of a general debiasing strategy, which comes from the non-visualization bias literature. Visualization debiasing strategies, which encode data specifically in a way to debias, or a combination of general and visualization debiasing strategies, may be useful in mitigating confirmation bias in non-experts. However, more work is needed to validate this.

VI. Acknowledgements

The authors wish to thank Qiyu Zhi and Eleanor Kinyon for assisting with experiment design. We also would like to thank Nicole McNeil, Jermaine Marshall, Maro Anuyah, Vivienne Armacost, Kaidan Fullerton, Kevin Wong, Annalisa Szymanski, Brianna Wimer, and Tariq Fuller for their feedback on the draft of this paper. Thanks to Natalie Whalen for assistance in data processing. This work is supported by the National Science Foundation under Grant No. 1816620

References

[1] A. Althubaiti. Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of multidisciplinary healthcare*, 9:211, 2016.

[2] D. Arnott. Cognitive biases and decision support systems development: a design science approach. *Information Systems Journal*, 16(1):55–78, 2006.

[3] E. Aronson. The theory of cognitive dissonance: A current perspective. *Advances in experimental social psychology*, 4(1):34, 1969.

[4] M. Card. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.

[5] S. Chaiken. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of personality and social psychology*, 39(5):752, 1980.

[6] S. Chaiken and Y. Trope. *Dual-process theories in social psychology*. Guilford Press, 1999.

[7] I. K. Choi, N. K. Raveendranath, J. Westerfield, and K. Reda. Visual (dis) confirmation: Validating models and hypotheses with visualizations. In *2019 23rd International Conference in Information Visualization–Part II*, pages 116–121. IEEE, 2019.

[8] M. B. Cook and H. S. Smallman. Human factors of the confirmation bias in intelligence analysis: Decision support from graphical evidence landscapes. *Human Factors*, 50(5):745–754, 2008.

[9] P. Croskerry. From mindless to mindful practice—cognitive bias and clinical decision making. *N Engl J Med*, 368(26):2445–2448, 2013.

[10] N. E. Dunbar, C. H. Miller, B. J. Adame, J. Elizondo, S. N. Wilson, B. L. Lane, A. A. Kauffman, E. Bessarabova, M. L. Jensen, S. K. Straub, et al. Implicit and explicit training in the mitigation of cognitive bias through the use of a serious game. *Computers in Human Behavior*, 37:307–318, 2014.

[11] P. Fischer, T. Greitemeyer, and D. Frey. Self-regulation and selective exposure: the impact of depleted self-regulation resources on confirmatory information processing. *Journal of personality and social psychology*, 94(3):382, 2008.

[12] D. Frey. Recent research on selective exposure to information. *Advances in experimental social psychology*, 19:41–80, 1986.

[13] W. J. Hall, M. V. Chapman, K. M. Lee, Y. M. Merino, T. W. Thomas, B. K. Payne, E. Eng, S. H. Day, and T. Coyne-Beasley. Implicit racial/ethnic bias among health care professionals and its influence on health care outcomes: a systematic review. *American journal of public health*, 105(12):e60–e76, 2015.

[14] C. Hill, A. Memon, and P. McGeorge. The role of confirmation bias in suspect interviews: A systematic evaluation. *Legal and Criminological Psychology*, 13(2):357–371, 2008.

[15] E. Jonas, S. Schulz-Hardt, D. Frey, and N. Thelen. Confirmation bias in sequential information search after preliminary decisions: an expansion of dissonance theoretical research on selective exposure to information. *Journal of personality and social psychology*, 80(4):557, 2001.

[16] J. Klayman and K. Brown. Debias the environment instead of the judge: An alternative approach to reducing error in diagnostic (and other) judgment. *Cognition*, 49(1-2):97–122, 1993.

[17] S. Knobloch-Westerwick, B. K. Johnson, and A. Westerwick. Confirmation bias in online searches: Impacts of selective exposure before an election on political attitude strength and shifts. *Journal of Computer-Mediated Communication*, 20(2):171–187, 2015.

[18] R. P. Larrick. Debiasing. *Blackwell handbook of judgment and decision making*, pages 316–338, 2004.

[20] J. Mackinlay. Automating the design of graphical presentations of relational information. *Acm Transactions On Graphics (Tog)*, 5(2):110–141, 1986.

[21] R. Mendel, E. Traut-Mattausch, E. Jonas, S. Leucht, J. M. Kane, K. Maino, W. Kissling, and J. Hamann. Confirmation bias: why psychiatrists stick to wrong preliminary diagnoses. *Psychological Medicine*, 41(12):2651–2659, 2011.

[22] R. S. Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2):175–220, 1998.

[23] J. Park, P. Konana, B. Gu, A. Kumar, and R. Raghunathan. Confirmation bias, overconfidence, and investment performance: Evidence from stock message boards. *McCombs Research Paper Series No. IROM-07-10*, 2010.

[24] P. Parsons. Promoting representational fluency for cognitive bias mitigation in information visualization. In *Cognitive Biases in Visualizations*, pages 137–147. Springer, 2018.

[25] J. Rajsic, D. E. Wilson, and J. Pratt. Confirmation bias in visual search. *Journal of experimental psychology: human perception and performance*, 41(5):1353, 2015.

[26] K. E. Stanovich. *Who is rational?: Studies of individual differences in reasoning*. Psychology Press, 1999.

[27] J. Steinberg. *The gatekeepers: Inside the admissions process of a premier college*. Penguin, 2003.

[28] P. T. Sukumar and R. Metoyer. A visualization approach to addressing reviewer bias in holistic college admissions. In *Cognitive Biases in Visualizations*, pages 161–175. Springer, 2018.

[29] P. Talkad Sukumar, R. Metoyer, and S. He. Making a pecan pie: Understanding and supporting the holistic review process in admissions. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–22, 2018.

[30] M. A. Tolcott, F. F. Marvin, and P. E. Lehner. Expert decision-making in evolving situations. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(3):606–615, 1989.

[31] F. Tschan, N. K. Semmer, A. Gurtner, L. Bizzari, M. Spychiger, M. Breuer, and S. U. Marsch. Explicit reasoning, confirmation bias, and illusory transactive memory: A simulation study of group medical decision making. *Small Group Research*, 40(3):271–300, 2009.

[32] P. C. Wason. On the failure to eliminate hypotheses in a conceptual task. *Quarterly journal of experimental psychology*, 12(3):129–140, 1960.