Confirmation Bias: The Double-Edged Sword of Data Facts in Visual Data Communication

Shiyao Li
Computer Science
Emory University
Atlanta, Georgia, USA
shiyao.li@emory.edu

Thomas James Davidson
Computer Science
Emory University
Atlanta, Georgia, USA
Thomas.James.Davidson@emory.edu

Cindy Xiong Bearfield
Interactive Computing
Georgia Institute of Technology
Atlanta, Georgia, USA
cxiong@gatech.edu

Emily Wall
Computer Science
Emory University
Atlanta, Georgia, USA
emily.wall@emory.edu

Abstract

Incorporating data facts, which are natural language descriptions of data patterns, alongside visualizations can guide readers and enhance the visibility of data patterns. However, data facts might also induce confirmation bias in visual analysis. We conducted a series of crowdsourced experiments to explore the biasing effects of data facts. Our findings show that the presentation style, strength, and alignment of data facts with pre-existing beliefs significantly impact confirmation bias. Data facts that support prior beliefs can exacerbate confirmation bias, whereas those that refute an individual's beliefs can mitigate it. This effect is amplified when data facts are used in combination with visual annotations. Data facts describing variable correlations are perceived to be more compelling than ones describing average values and are associated with higher levels of confirmation bias. We underscore the persuasive influence of data facts in visualizations and caution against their indiscriminate use in efforts to mitigate bias.

CCS Concepts

• **Human-centered computing → Empirical studies in visualization**.

Keywords

confirmation bias, data visualization, data facts, annotation

ACM Reference Format:
Shiyao Li, Thomas James Davidson, Cindy Xiong Bearfield, and Emily Wall. 2025. Confirmation Bias: The Double-Edged Sword of Data Facts in Visual Data Communication. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3706598.3713831

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713831

1 Introduction

People can easily become overwhelmed by data and exhibit biased behavior in visual analysis and decision-making [17, 63, 64]. One particularly insidious bias in this context is confirmation bias––the tendency to seek out and interpret information in ways that affirm preexisting beliefs [43]. In data communication and visual analytics, this bias is especially problematic when a visualization allows for multiple interpretations, as it shapes the takeaway to align with prior beliefs, thereby hindering objective data analysis.

A common misconception among visualization readers is that visualizations present objective truths, leading them to believe that 'what they see' is inherently 'what is.' This assumption can cause even seasoned analysts to fall prey to confirmation bias, seeing only what aligns with their beliefs and expectations and, as a result, drawing inaccurate or sub-optimal conclusions. Research in psychology and visualization has shown that confirmation bias is notoriously difficult to overcome [73]. Even awareness of the bias offers little protection [17, 29].

This phenomenon can be further complicated by the complex interactions between visualizations and text. Textual annotations, which we define as text-based notes added to explain or comment on visual patterns in visualizations, can emphasize key messages to help readers spot relevant data patterns efficiently [6, 28, 58] and even shape the interpretation of the data [30, 31, 57]. In this work, we investigate the impact of a specific type of textual annotation (data facts) on confirmation bias in visual data communication. **Data facts** are natural language descriptions of data insights ranging from pointing out a single value in a visualization to complex comparisons of data distributions [56].

Widely used visualization tools, such as Tableau, Microsoft Power BI and Google Charts [1, 14, 24], often leverage the power of automation to generate data facts from visualizations. The intent is to help data analysts extract data patterns more efficiently [25]. However, these data facts may subtly bias people [57]. For instance, consider data facts in the form of searchable natural language trends in the data (e.g., the Voder system [56]). An analyst might cherry-pick evidence that supports their hypothesis while ignoring contradictory data. Alternatively, data facts might encourage critical thinking, offering the analyst alternative perspectives that challenge pre-existing beliefs and reveal overlooked patterns. Thus in this work, *we examine the potential of data facts to mitigate or exacerbate confirmation bias in visual analysis.* We empirically test these intuitions to systematically assess the risks and benefits of displaying data facts with visualizations.

1.1 Research Questions and Hypotheses

Across three preliminary studies and one primary experiment, we ask the research question:

RQ: How does the *presence*, *presentation style*, and *strength* of data facts impact the manifestation of confirmation bias in visual data communication?

We explore how confirmation bias may be exacerbated or mitigated by data facts (in our experiment design, we operationally refer to data facts as brief textual summaries highlighting statistical patterns in visualizations) when combined with visualizations across five hypotheses. To better understand the factors influencing confirmation bias in visual data communication, we draw on a range of psychological and data visualization research to formulate hypotheses.

First, we investigate the effect of **prior belief** on confirmation bias. Previous research in psychology has shown that confirmation bias can fluctuate in severity depending on the relationship between the user's stance and the displayed information. For example, confirmation bias is more pronounced when people hold stronger attitudes [7]. Hence, we examine whether the impact of prior belief extends to the interpretation of visual communication. This will also help us account for the effect of prior belief in later analyses to reach more generalizable conclusions on the effect of other experimental factors.

H1 - Prior Belief Strength: Participants with *stronger* prior beliefs will exhibit *stronger* confirmation bias.

Similar to prior beliefs, the data topic is another contextual factor that is difficult to control when visualization researchers build tools or generate best-practice guidelines for visual analysis in the real world. Data topics might interact with the strength of one's prior beliefs to consequently impact the degree of confirmation bias. For instance, political or polarizing topics in which people tend to have stronger prior beliefs may be more likely to exacerbate confirmation bias than more neutral topics [16, 62, 71].

H2 - Topic: Confirmation bias will be *higher* with a more *polarizing* dataset (COVID-19 Vaccination) compared to a less polarizing dataset (Diet Choices).

With the previous two hypotheses and subsequent analysis, we can build models to control for the effect of prior belief and data topics to investigate the role of data facts in influencing confirmation bias. Data facts can either ***support*** or ***refute*** one's prior beliefs to differently affect confirmation bias. Prior research has revealed that individuals tend to perceive a visualization as less credible when the title is ***inconsistent*** with their prior beliefs compared to when the title is ***consistent*** with their beliefs [30]. We thus test how **data facts and belief alignment** impacts the manifestation of confirmation bias.

H3 - Data Facts Alignment: Confirmation bias will be higher when visualizations are accompanied by data facts that ***support*** participants' prior beliefs, and lower when the visualizations are accompanied by data facts that ***refute*** participants' prior beliefs, compared to when the visualizations are shown without any data facts.

We also investigate the **presentation style** of data facts. More visualizations are generated in analytic environments with accompanying data facts [56, 69]. Further, textual data facts are often accompanied by visual annotations (in our experiment design, we operationally define visual annotations as graphical elements that emphasize a specific data point or pattern) to enhance the effectiveness of data presentation [10, 51, 52, 55]. Recent advances have even integrated auto-generated visual annotations, derived from textual descriptions of data, to streamline the annotation process [33]. These advancements in technology motivate us to investigate how representing data facts with visual annotations and textual annotations affect confirmation bias.

H4 - Data Facts Presentation Style: Confirmation bias will be strongest for participants who view data visualization with ***supporting*** textual data facts only, followed by textual data facts with visual annotations, and then ***No Data Facts***. For ***refuting*** data facts, confirmation bias will be lower when participants view visualizations with textual data facts only, followed by textual data facts with visual annotations.

Finally, we consider the **strength** of the data facts themselves. Previous studies have demonstrated that captions highlighting *key visual features* in data visualizations make these features more salient [74]. Readers often interpret these captions as the primary chart takeaway [28]. Building on this, we examine whether this effect extends to data facts. We generated a set of data facts that describe different insights, referencing the taxonomies proposed by Amar et al. [2], to create a ranked list of data facts varying in persuasive **strength**. We test the effect of data facts on confirmation bias based on their strength.

H5 - Data Facts Strength: Confirmation bias will be higher when ambiguous visualizations are presented with ***supporting*** *stronger* data facts (i.e., describing correlations) compared to ***supporting*** *weaker* data facts (i.e., describing a derived value). Confirmation bias will be lower when visualizations are presented with ***refuting*** correlation data facts compared to ***refuting*** derived value data facts.

1.2 Summary of Findings

A summary of our findings can be found in Table 1. Our results showed that data facts can *exacerbate* confirmation bias when their phrasing ***supports*** an individual's prior beliefs, and this effect is further exacerbated when accompanied by visual annotations. Conversely, data facts can *mitigate* confirmation bias when their phrasing ***refutes*** an individual's prior beliefs. Our findings highlight a critical drawback of incorporating textual summaries into data visualizations: data facts can introduce bias, particularly when they align with individuals' prior beliefs. However, we also found that data facts show promise in bias mitigation when the data fact supports a refuting hypothesis. We discuss the implications of these findings for visualization design and suggest that careful consideration is needed when adding textual annotations—whether as titles or descriptions—to ensure that the interpretation remains balanced and not overly influenced by these textual elements.

Hypothesis		Support	Finding Summary
H1	Participants with *stronger* prior beliefs will exhibit *stronger* confirmation bias.	Supported	The strength of pre-existing beliefs is positively correlated with the level of exhibited confirmation bias.
H2	Confirmation bias will be *higher* with the more *polarizing* dataset (COVID-19 Vaccination) compared to the less polarizing dataset (Diet Choices).	Not Supported	The polarization level of the visualized topic does not have an effect on confirmation bias.
H3	Confirmation bias will be higher when visualizations are accompanied by data facts that ***support*** participants' prior beliefs, and lower when the visualizations are accompanied by data facts that ***refute*** participants' prior beliefs, compared to when the visualizations are shown without any data facts.	Supported	***Supporting*** Data Facts lead to a higher level of confirmation bias compared to when there are ***No Data Facts***.
			Refuting Data Facts result in a lower level of confirmation bias compared to when there are ***No Data Facts.***
H4	Confirmation bias will be strongest for participants who view data visualization with ***supporting*** textual data facts only, followed by textual data facts with visual annotations, and then ***No Data Facts***. For ***refuting*** data facts, confirmation bias will be lower when participants view visualizations with textual data facts only, followed by textual data facts with visual annotations.	Partially Supported	***Supporting*** Data Facts with Visual Annotation lead to a greater level of confirmation bias compared to when there are ***No Data Facts***.
			No significant difference was identified between ***Refuting*** Data Facts with Visual Annotation, ***Refuting*** Data Facts only, and ***No Data Facts***.
H5	Confirmation bias will be higher when ambiguous visualizations are presented with ***supporting*** strong data facts (i.e., describing correlations) compared to ***supporting*** less strong data facts (i.e., describing a derived value). Confirmation bias will be lower when visualizations are presented with ***refuting*** correlation data facts compared to ***refuting*** derived value data facts.	Partially Supported	***Supporting*** Correlation Data Facts lead to a higher level of confirmation bias compared to ***No Data Facts***.
			Refuting Correlation Data Facts result in a lower level of confirmation bias compared to ***No Data Facts***.

Table 1: Results on Research Question Hypotheses. Data facts that *support* prior beliefs tend to intensify confirmation bias, while those that *refute* prior beliefs can mitigate it. Data facts describing key visual features can further amplify this effect. Additionally, visual annotations with *supporting* data facts exacerbate confirmation bias. Rows highlighted in blue indicate hypotheses that received partial or full support.

2 Related Work

2.1 Data Facts

Natural language descriptions of data patterns, so-called data facts, are a common feature in many visualization systems such as Voder [56], PowerBI QuickInsights [18], and others [34, 37, 39, 45, 50, 58]. They are often used with charts to summarize or highlight key insights [4, 56], or to guide further analysis by recommending alternate visualizations to explore [18, 56]. Use of data facts for data communication and exploratory data analysis offers several advantages including reducing manual labor involved in data exploration [4, 35], increasing reader engagement, interpretation& sense-making[4, 34, 50, 58], and improving accessibility for individuals with low vision and/or low data literacy [45, 58]. While researchers acknowledge challenges due to the use of automated generated data facts in general – such as user mistrust [35, 37, 56], spurious insights [4, 13] and the potential for increase in cognitive biases [37, 49] – there is little empirical evidence to support these concerns related to data facts in particular.

2.2 Impact of Textual Content Alongside Data Visualization

Prior studies from Stokes et al. have demonstrated that readers exhibit a strong preference for integrating textual content and data visualizations for enhanced comprehension [58]. Meanwhile, a combination of text content (such as titles, captions) in addition to visualizations can significantly impact information recall, reasoning, and decision-making processes of the readers. For example, earlier work has shown that visualization readers are more likely to recall information conveyed by a biased title than the actual graphic contents [30, 31]. Captions that match the key visual feature in data visualizations enhance the likelihood of viewers perceiving this feature as the main takeaway while mismatches between captions and charts can lead to overlooked details from the captions [28].

Additionally, researchers have explored how the interaction between accompanying text and participants' pre-existing beliefs shapes individuals' perceptions of data visualization. Kong et al. observed individuals perceiving a data visualization as less credible when its title was inconsistent with their pre-existing beliefs [31], even when the data visualization contents remained consistent with those beliefs. More recently, Stokes et al. [57] found that people perceive the authors of a visualization to be more biased if the visualization is accompanied by textual annotations. The amount of perceived bias is exacerbated when the text annotation conflicts with the readers' pre-existing beliefs. We focus on confirmation bias and build on prior work in two key ways. First, we move beyond textual annotations in visualizations to also examine the impact of visual annotations. Second, we shift from evaluating perceived biases in the author's stance and credibility to quantifying how strongly a visualization supports or refutes a particular stance. These manipulations enable us to derive quantitative models that predict how effectively a visual or textual annotation can nudge individuals toward a certain interpretation, given their prior beliefs and the visual evidence presented.

2.3 Confirmation Bias

In psychology, confirmation bias is a widely recognized inferential error where individuals seek and interpret evidence that aligns with their pre-existing beliefs, expectations, and hypotheses [21, 43]. This tendency has been extensively investigated in other disciplines as well. For instance, sociological studies reveal that people consistently engage with news that supports their pre-existing beliefs [38], while scholars often favor evidence supporting their preferred theories without thoroughly considering contradicting perspectives [54]. In political science, Taber and Lodge highlighted how the tendency to seek confirmatory evidence among individuals with conflicting political stances significantly contributes to polarization in political beliefs [59].

In the context of data visualizations, prior research has developed methodologies to detect and counteract confirmation bias [11, 15, 42, 44, 75]. This bias manifests when individuals interpret visualizations in a way that aligns with their prior beliefs, leading them to overemphasize supportive visual patterns while downplaying contradictory information. Additionally, researchers have found that the clarity and complexity of visualizations can significantly influence their perceived trustworthiness [19, 48]. Psychological studies further show that confirmation bias is more pronounced among individuals with strongly held pre-existing beliefs [7], suggesting that both the strength of pre-existing beliefs and the perceived reliability of visual evidence play crucial roles in the manifestation of this bias.

Motivated by these insights, our experiments aim to assess the *magnitude* of confirmation bias when individuals reason with data visualizations accompanied by textual annotation. Our research goes beyond merely detecting the presence of confirmation bias to understanding how these elements interact to influence the extent of bias in interpretation.

3 Study Overview

In this work, we examine the effect of visual and textual annotation on confirmation bias in visualizations. This type of investigation is sensitive to contextual factors such as the visualization topic, the complexity of the data visualization, the patterns in a visualization that the data facts describe, and our participants' prior beliefs. Therefore we conducted a series of preliminary studies (PS) to identify the optimal experimental parameters for our main study, including:

- *PS 1*: a range of topics with various prior belief distributions among participants,
- *PS 2*: visualization stimuli that are *ambiguous* and, according to an individual's prior beliefs, could be selectively interpreted as supporting or refuting, and
- *PS 3*: a set of data facts with different strengths to explore their impact on confirmation bias.

We conducted our main study using experimental parameters determined by the preliminary study. We provide a list of the dependent variable (DV) and independent variables (IV) we examined for the impact on confirmation bias in our main study.

- *DV*: amount of confirmation bias
- *IV 1 (H1)*: prior belief strength
- *IV 2 (H2)*: chart topics (vaccination, dietary choices)
- *IV 3 (H3)*: data facts alignment (supporting, refuting)
- *IV 4 (H4)*: data facts presentation style (none, textual, textual + visual)
- *IV 5 (H5)*: data facts strength (correlation, derived value)

4 Preliminary Studies

Participants for all preliminary studies were recruited through the crowdsourcing platform Prolific [46]. We screened for participants who are based in the United States who are fluent in English and have a study approval rate greater than 95%. Individuals who participated in any one study were ineligible to participate in subsequent ones. The study plans were reviewed and approved by the authors' university ethics board.

4.1 Preliminary Study 1: Data Topic Selection

Existing beliefs significantly impact how one processes information. The topic of a visualization can trigger varying beliefs, making the reader more or less susceptible to confirmation bias [26]. Topics unfamiliar to readers are less likely to elicit confirmation bias due to a lack of pre-existing beliefs, whereas polarizing or political topics are more likely to elicit confirmation bias as individuals' entrenched stances are less likely to move [40]. To enhance the robustness of our work, we tested various topics to identify ones associated with strong and weak prior beliefs. The main study uses topics selected from this set. This approach allows us to account for the effects of prior belief strength.

Participants. We recruited 40 participants from the crowdsourcing platform Prolific (M_{age} = 33.59, SD_{age} = 12.47; 20 Male, 18 Female, 1 Non-binary, 1 Prefer not to disclose).

Tasks & Procedure. Participants were presented statements on 16 topics (e.g., vaccine, dietary preference, government surveillance) in a random order, such as "COVID vaccination should be a mandatory policy." The full list can be found in Figure 1. Following a similar set-up as prior work that used belief elicitation (e.g., [73]), we asked participants to rate the "extent to which they agree with the statement" on a scale ranging from "-100 = Strongly Disagree" to "0 = Neutral" to "100 = Strongly Agree." The study ended with several optional demographic questions including their age and gender.

Measuring Topic Polarization. The data we collected presents a distribution of scores for each topic. We employed a continuous slider labeled with "Strongly Disagree" on the left, "Neutral" in the center, and "Strongly Agree" on the right. The numerical range of the slider was not visible to participants, who only saw the labels on the slider ("Strongly Disagree", "Neutral", and "Strongly Agree"). We categorized scores below 0 into the 'Disagree' cluster, and those exceeding 0 in the 'Agree' cluster. We aimed to select a topic associated with strong, polarizing beliefs (so that regardless of participants' stances the topic would elicit strong beliefs), as well as a neutral topic associated with weak beliefs. To this end, we computed the Cluster-Polarization Coefficient (CPC) [41] to measure the polarization of each topic. CPC captures two important features of political and social polarization: intergroup heterogeneity (*BSS*, which represents the variance between each cluster) and intragroup homogeneity (WSS, which represents the sum of variance within all clusters). Subsequently, we employed Equation 1 to compute the CPC score. The score indicates the polarization level for each topic, with 0 denoting no polarization (e.g., all participants consistently adjusted the slider toward either the left (disagree) or right (agree) side of the central point "Neutral," indicating unanimous disagreement or agreement with the statement) and 1 indicating extreme polarization (e.g., all participants in the disagree cluster moved the slider to the exact same position at the left side of the central point, while all participants in the agree cluster moved it to the exact same position at the right side of the central point). Responses with neutral beliefs were excluded when calculating CPC (3.4% of total responses), as our goal was to explore polarization on both the 'Disagree' and 'Agree' sides of each topic.

$$CPC = 1 - \frac{WSS}{WSS + BSS} = \frac{BSS}{WSS + BSS} \quad (1)$$

Results. We ranked all topics based on the computed CPC scores (full details in supplementary materials). One potential limitation of the CPC metric in our experiment setting could be that a CPC of 1 indicates extreme polarization, but does not necessarily reflect strongly held beliefs, as it could also result from participants agreeing and disagreeing uniformly on weak points. Consequently, we validated our topic selection by examining the belief distributions (as shown in Figure 1) to ensure the selected topics represent both the desired polarization *and* strength of belief. The topics we ended up selecting include:

- High polarization topic associated with stronger beliefs: *COVID-19 Vaccination* (CPC: 0.74).
  We selected 'Vaccination' as the polarizing topic based on its highest CPC score and the concentration of strong beliefs within both the agree and disagree clusters.
- Neutral topic associated with weaker beliefs: *Diet Choices* (CPC: 0.65).
  We selected 'Diet' as the neutral topic because it has the highest percentage of beliefs near the neutral point (0), with 15% of participants within the [-5, 5] range. Although two other topics, 'Immigration' (CPC = 0.61) and 'Work from Home' (CPC = 0.51), also have 15% of beliefs near neutral, 'Diet' maintains a higher level of polarization (0.65). We chose the topic with the higher CPC score because it reduces the likelihood of a skewed belief distribution (e.g., topic 16-Pet in Figure 1) among participants, which could hinder their reception to evidence that contradicts participants' common assumptions.

4.2 Preliminary Study 2: Visualization Stimuli Selection

Existing work has operationalized confirmation bias as the tendency to interpret ambiguous evidence in a way that supports one's prior beliefs [9, 12, 29]. Building on this idea, we measure confirmation bias by examining how much participants perceive an ambiguous data visualization—one that neither strongly supports nor refutes an existing belief—as supporting their own beliefs. A visualization can contain multiple patterns. Readers can draw different inferences about data depending on what patterns they pay attention to [6]. To conduct our main study, we need a set of ambiguous visual stimuli that are just complex enough to allow for various interpretations but still straightforward for participants to engage with. This preliminary study tackles this goal via two stages. First, we identify the optimal data set size so the visualization is complex but not overwhelming. Second, we tweak the data values to generate ambiguous patterns in data that can be interpreted in competing directions. Our goal is to create stimuli with evenly divided interpretations: approximately 50% of the responses should indicate seeing a predominantly increasing trend, and the rest a decreasing trend. Using ambiguous visualizations for our study not only provide us the max amount of space to observe variance in data patterns to avoid floor and ceiling effects, but also provides a clean, conservative measure of confirmation bias as the observed confirmation bias will be less influenced by the visualization's inherent narrative. If we instead used a non-ambiguous visualization that explicitly takes a stance, we would need to model its relative position to each participant's belief, introducing additional complexity and noise.

4.2.1 Stage 1: Determining Small Multiples Size.

Materials: Prior work suggested that bar charts depicting groups of data can elicit ambiguous takeaways [57]. Therefore we generated four types of faceted bar charts: arranged in either 3×3 or 5×5 grids, with each facet cell within these grids featuring either a single bar or two bars representing different groups. To introduce ambiguity, we balanced the number of rows displaying strictly ascending and descending trends in bar heights. We also incorporated rows that presented neutral trends (an equal number of descending and ascending trends between adjacent cells). For each design, we created two datasets to further examine whether the ambiguity persisted across different data sets. All stimuli from this stage are included in the supplemental materials.

Participants: We recruited 83 participants (M_{age} = 39.23, SD_{age} = 13.84; 44 Male, 39 Female) from Prolific. Each participant was randomly assigned to one of our visualization stimuli (3x3 single-bar, 3x3 multi-bar, 5x5 single-bar, or 5x5 multi-bar).

Study Procedure: We designed a visual search task to evaluate the complexity of each visualization following guidelines outlined in Elliott et al. [20]. Each participant went through three visual search tasks in a random order. Each task involved presenting a visualization and asking users to determine whether it displayed an increasing or decreasing trend in row-wise bar heights using a two-choice forced response. We counterbalanced the order of the response options (descending first or ascending first) to prevent an ordering effect. Two of the tasks used the faceted bar charts we designed. The third task served as a comprehension check, using a visualization that depicted a strict descending trend across all rows, to assess the participants' ability to extract trends from visualizations. To avoid the impact of participants' pre-existing beliefs on data interpretation, we created hypothetical scenarios about alien fictional gases (Gas Task) and chemicals in lake (Chemical Task) for this study. We provide one example task below. The rest can be found in the supplementary materials.

- *Example Task (Gases Task)*: You need to determine which of two fictional gases, Zytharane or Skylox, is dominant in a

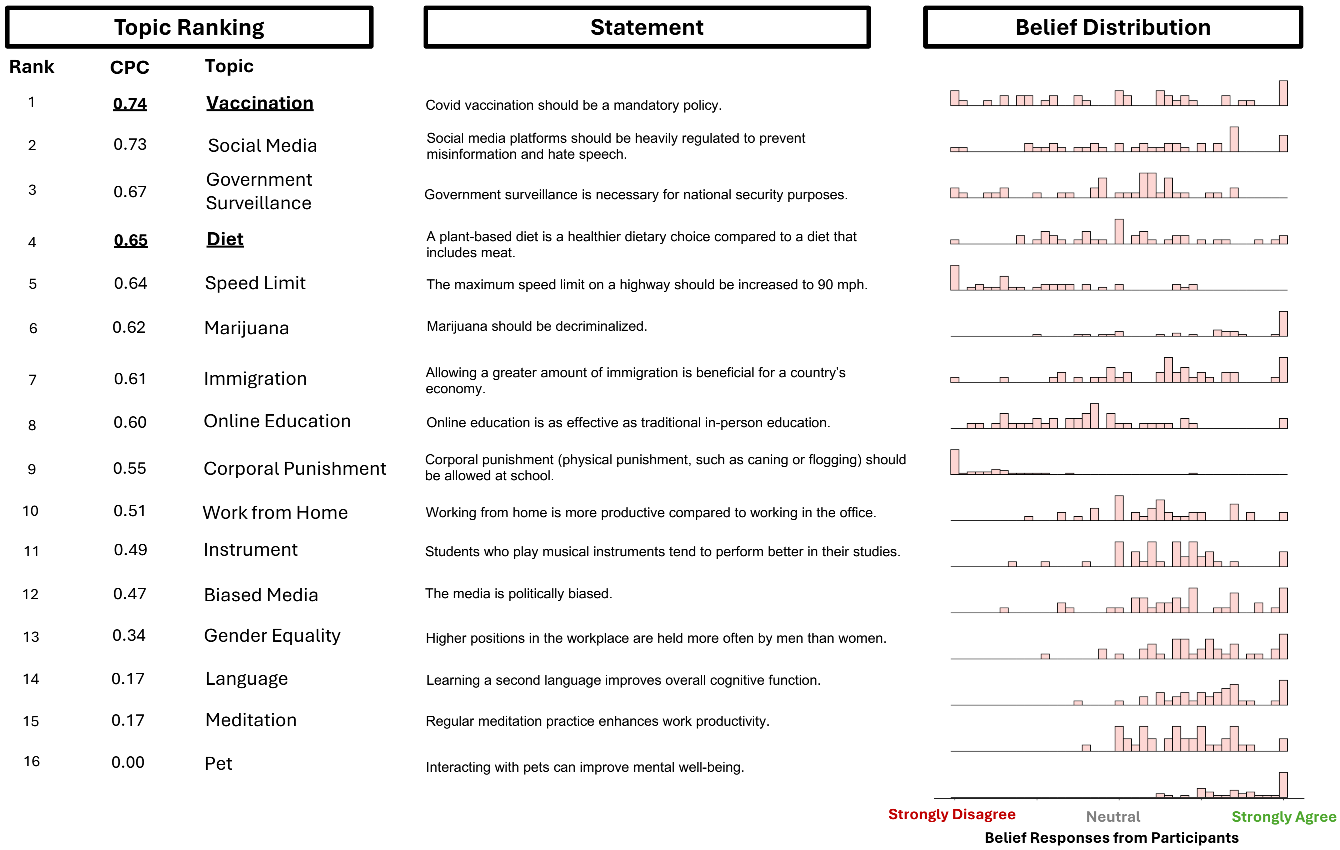

Figure 1: Preliminary Study 1 Results and Belief Distributions. Participant responses were collected as slider inputs and visualized as histograms for each topic, with a bin width of 5 for each bar. The x-axis ranges from -100 to 100, with -100 representing 'Strongly Disagree,' 0 representing 'Neutral,' and 100 representing 'Strongly Agree.' We selected 'Vaccination' as the polarizing topic, as individuals tend to hold strong prior beliefs about it, and 'Diet' as the neutral topic, where people generally hold weaker beliefs.

> sample consisting exclusively of these gases. Zytharane is characterized by a warming effect, where its concentration increases with rising temperatures. Conversely, Skylox exhibits a cooling effect, with its concentration decreasing as temperatures rise. This data visualization depicts the relationship between the atmospheric concentration of the gas mixture and temperature changes. To examine the dominant gas in mixture S, please analyze whether the visualization shows an overall ascending or descending trend in the atmospheric concentration as the temperature increases.

Each participant was forced to wait a minimum of 20 seconds before answering to encourage a proper examination of each visualization. After completing the three tasks, they were asked demographic questions including age, gender, and the highest level of education.

Results: As shown in Figure 2 A , the 5×5 multi-bar condition led to the most even split in participant response. We took a closer look at participant feedback from Prolific direct messages. Many indicated that the multi-bar visualization was confusing, and as a result, they incorrectly approached the task by comparing the two bars within the same cell for ascending or descending trends, rather than comparing the overall bar heights across the same row. Further, the 3×3 seemed was too simplistic of a visualization, such that participants overall saw either only the ascending or the descending trend (depending on the dataset). Therefore, we decided to go with the 5×5 single-bar visualization.

4.2.2 Stage 2: Generating Ambiguous Data. In order to further improve our 5×5 single-bar visualization to be consistently ambiguous, we adjusted the specific data values to reach an even 50-50 split amongst participants in terms of reported ascending or descending trends.

Materials. We developed a Python script that simulates data values in a 5×5 matrix to test a range of stimuli. The script enforces the balance of the following patterns for each row to maintain ambiguity (see supplemental materials for more details):

- **Pattern 1 Strictly Ascending/Descending:** The heights of the five bars in a row progressively increase or decrease from left to right.

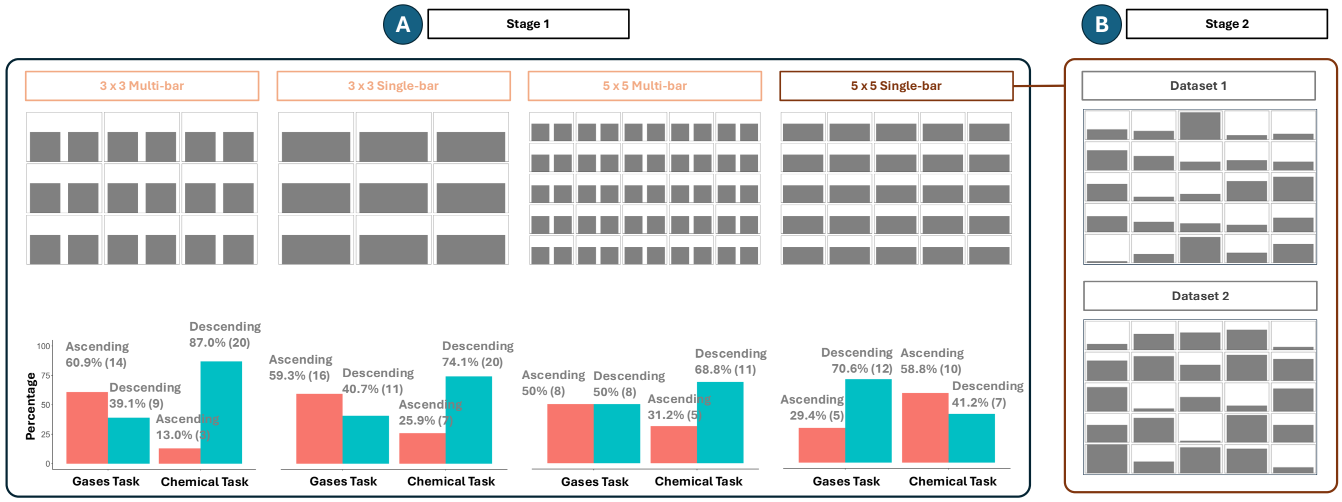

Figure 2: Preliminary Study 2: Stage 1 (A) and Stage 2 (B) Conditions. In Figure 2 A, the depicted charts under assessed dimensions are simplified graphical representations of the stimuli, the actual depicted charts in our study include titles, axis labels, and tick marks (see supplemental materials). At the bottom of Figure 2 A, the bar charts illustrate the distribution of evaluations for each stimulus dimension in Stage 1. We selected 5×5 single-bar stimuli, colored in brown, and chose datasets 1 and 2 from the 16 generated datasets in Stage 2 as depicted in In Figure 2 B.

- **Pattern 2 Overall Ascending/Descending:** For a given row, four out of five bars follow a strict ascending or descending sequence from left to right.
- **Pattern 3 Neutral trend:** For a given row, the height difference between any symmetrical pair of bars (e.g., the first bar on the left and the last bar on the right, the second and the fourth bars) will not exceed 3 units. This specification keeps enough variability in the neutral trend to appear realistic.

To ensure the ambiguity of each resulting visualization, the script ensures there exists an equal number of rows following Patterns 1 and 2. We simulated 16 ambiguous data visualizations that satisfy our criterion of ambiguity.

Participants. We recruited 100 participants (M_{age} = 37.92, SD_{age} = 12.38, 45 Male, 53 Female, 1 Non-binary, 1 Prefer not to disclose) from Prolific and conducted two rounds of this preliminary study (50 participants, 8 visualizations in each round).

Study Procedure. In each round, participants examined eight ambiguous visualizations in random order and answered the question, *"Based on the visualization provided, please identify the general trend in bar heights as you move from left to right across each row."* Participants then selected one of two options: Ascending or Descending. To ensure a thorough evaluation, each participant was required to wait a minimum of 20 seconds before responding. We also randomly inserted a control visualization with five strictly ascending rows as a comprehension check to assess each participant's ability to reason whether the majority of rows depict an overall ascending or descending trend. Participants who failed the comprehension check were excluded from the data analysis.

Results. An ideal ambiguous visualization would be one in which participants responded in equal proportions that the visualization depicted an ascending v. descending trend (i.e., 50% ascending, 50% descending). To account for noise, we consider stimuli sufficiently ambiguous if participants' response distributions fell between 40% and 60% ascending/descending. Five stimuli satisfied this condition. Among them, we selected the two visualizations that elicited responses closest to 50% ascending/descending, as depicted in Figure 2 B.

4.3 Preliminary Study 3: Strength of Data Facts

Data facts can vary in their persuasive strength [47, 57]. To account for the effect of data fact strength on confirmation bias, we conducted this study to identify data facts with strong and weak persuasiveness.

Participants. We recruited 40 participants (M_{age} = 33.35, SD_{age} = 9.29; 21 Male, 17 Female, 1 Non-binary, 1 Prefer not to disclose) from Prolific. Each participant read visualizations with ten types of data facts in random order.

Materials.
<u>Data Facts:</u> We designed ten types of data facts aligned with the ten low-level analytic tasks delineated in Amar et al.'s framework, similar to the approaches by Srinivasan et al. [2, 56]. These examples are included in the supplemental materials.
<u>Visualization:</u> We used one of the 5×5 small multiples bar charts as depicted in Figure 2, depicting the vaccination topic. This visual stimulus is also included in the "Example Visual Stimuli and Data Facts" panel in Figure 3.

Study Procedure. After obtaining informed consent, each participant read 10 visualizations, each accompanied by a line of text describing a unique data fact, in random order. We employed the same slider settings as described in Section 4.1 to elicit participants' prior beliefs on vaccinations with the prompt: "Please rate the extent

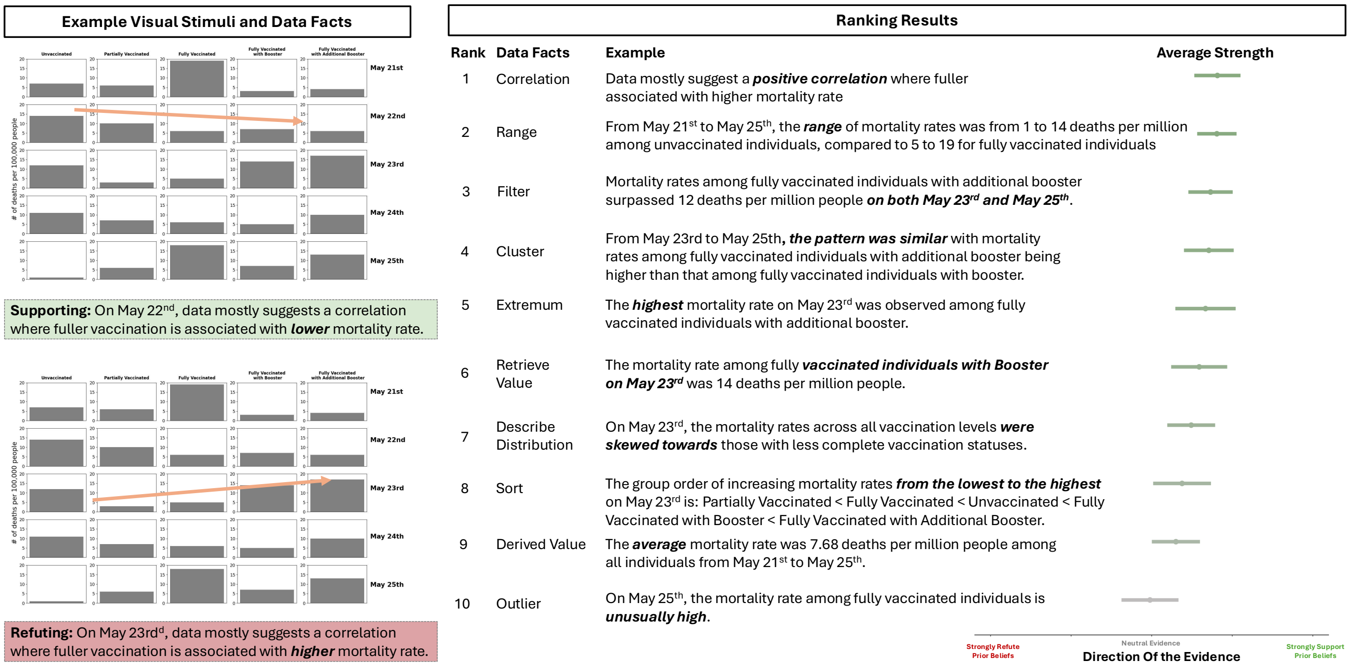

Figure 3: Preliminary Study 3: Data Facts and Results. The strength of each type of data fact was assessed based on how strongly participants perceived the data visualizations and facts as supporting their prior beliefs. The mean value and 95% confidence intervals for each type of data fact are depicted in the confidence interval plot. The color coding indicates whether the mean rating of each data fact leans toward supporting or refuting participants' prior beliefs.

to which you agree with the following statement: COVID vaccination should be a mandatory policy." To measure the *confirmatory* effect of each type of data fact, we manipulated all data facts to align with participants' prior beliefs. We also included an attention check question: *"If you are paying attention, please drag the handle to the left end of the slider. This is an attention check."* For each visualization, participants evaluate it by responding to the prompt: *"How strongly do you perceive this visualization and the accompanying text to* ***support*** *the takeaway: COVID vaccination should be a mandatory policy?"* Participants responded using an interactive slider ranging from strongly refuting (-100) to strongly supporting (100). The numerical ranges for both sliders, eliciting prior beliefs and visualization evaluations, were hidden from participants.

We also included a comprehension check at the beginning of the study to ensure participants understood how to read our visualization stimuli. We asked: *"What was the mortality rate on May 22nd among partially vaccinated individuals?"* as a multiple-choice question. Participants who failed the attention check or the comprehension check were excluded from the analysis. Participants who held a neutral belief were also excluded as this makes it impossible to interpret potentially confirmatory belief updating.

Results. For each data fact, we calculated its perceived data fact strength score following Equation 2. The term *Response* represents the input from each participant's interactive slider. The data fact strength score reflects the extent to which the participant perceived the visualization (which was designed to be ambiguous and neutral) to be aligned with their prior beliefs. We designed the experiment such that participants only interacted with visualizations containing supportive data facts for evaluating confirmatory effects. A positive data fact strength score (+) means the participant reported the visualization to align with their prior beliefs, suggesting that the data fact biased the visualization toward those beliefs. Conversely, a negative score (-) means the participant reported the visualization to refute their beliefs, indicating that the data fact biased the visualization against those beliefs despite the data fact being designed to be supportive.

$$Data\ Facts\ Strength = \pm\,|response| \tag{2}$$

We calculated the average strength score for each type of data fact across all participants to produce a ranking of the strength for each data fact, as summarized in Figure 3. We ultimately selected 'Correlation' as the strong data fact and 'Derived Value' as the weak data fact for our main study. We used 'Derived Value' instead of 'Outlier' as the data facts with the lowest strength because the 'Outlier' condition resulted in a negative score, which could suggest unfamiliarity with the concept and potential misunderstanding. Hence we avoided using this data fact type.

5 Main Study

After the three preliminary studies, we conducted our main study to address our research questions in a pre-registered experiment[1].

[1] https://aspredicted.org/N78_C7G

Table 2: The statement for each topic

Topic	Statement
COVID-19 Vaccination	COVID vaccination should be a mandatory policy.
Diet Choices	A plant-based diet is a healthier dietary choice compared to a diet that includes meat.

Figure 4 shows the experimental variables and procedure. A summary of our findings is presented in Table 1, with more detailed analyses available in the supplemental materials.

Materials.

- *Topics*: To validate the impact of data facts across a broader range of pre-existing belief strengths, we selected two topics as determined from the Preliminary Study 1:
  - **[COVID-19 Vaccination]** A more polarizing topic where individuals tend to hold strong beliefs either in support of or opposition to mandatory COVID-19 vaccination.
  - **[Diet Choice]** A less polarizing topic where individuals tend to have relatively more neutral beliefs on whether a plant-based diet is healthier than one that includes meat.
- *Visual Stimuli*: Based on the results of Preliminary Study 2, we selected the ambiguous 5×5 small multiples bar charts with a single bar per cell as the stimuli, with two synthetic datasets.
- *Data Facts*: Two types of data facts with varying strengths accompanies our visualizations, based on the results of Preliminary Study 3.
  - **[Correlation]** A more compelling data fact describing the correlation between column and row variables.
  - **[Derived Value]** A less compelling data fact describing the average value of certain columns.

Conditions. Participants were randomly assigned to one of three presentation styles as between-subjects conditions: textual data facts with visual annotations, textual data facts only, or ***No Data Facts***. In all conditions, each participant read two visualizations, with the topic and datasets counterbalanced with the display order, as shown in Figure 4. For those in the textual data facts only and the textual data facts with visual annotations conditions, each participant read both a visualization with a correlation data fact and one with a derived value data fact, one of which supports their prior belief and the other refuting. Their display order and combination are counterbalanced. For example, if the first visualization showed a supporting correlation data fact on the topic of vaccines via dataset 1, the second visualization showed a refuting derived value data fact on the topic of diet choices via dataset 2. For the ***No Data Facts*** condition, participants viewed the two ambiguous visualizations without any data facts, with the dataset topic order randomized.

Task and Procedure. Our study procedure is shown in Figure 4. After completing the consent form, participants start with a warm-up activity to familiarize them with our stimuli. During the warm-up, they read chart descriptions explaining information encoded by rows and columns, followed by a comprehension check to ensure they understood how to interpret the chart. Participants who do not pass this check are excluded from our analysis.

Next, we assessed participants' pre-existing beliefs about each topic using an interactive slider ranging from "strongly disagree" (-100) to "strongly agree" (100) with the prompt: "Please rate the extent to which you agree with the following statement:" Statements related to the topics were provided as outlined in Table 2. Participants then read the bar charts, either with a textual data fact, with a textual data fact plus visual annotations, or with ***No Data Facts***. Afterward, they were prompted to rate their level of agreement with the same statements in Table 2, using an interactive slider ranging from "strongly ***refute***" (-100) to "strongly ***support***" (100). The numerical ranges of both sliders were not visible to participants. An additional trial on the topic "Productivity while working from home" was included as an attention check and appeared randomly between the two main trials: *"If you are paying attention, please drag the handle to the left end of the slider. This is an attention check."* Participants who failed this check were excluded from our analysis. Participants who held a neutral belief were also excluded as this makes it impossible to interpret potential confirmation bias.

Measuring the Strength of Prior Beliefs: We measure the strength of each participant's prior belief on each topic as the absolute distance from either endpoint of the slider ("strongly disagree" and "strongly agree"), based on the following equation:

$$belief\ strength = |initial\ belief| \tag{3}$$

In Equation 3, $initial\ belief$ represents the participant's response from the interactive slider. If a participant ***agrees*** with the statement provided, we measure the strength as the absolute distance from the right end of the slider. Conversely, if a participant ***disagrees*** with the statement, we measure the strength as the absolute distance from the left end of the slider.

Measuring Confirmation Bias. We evaluated how participants interpreted an ambiguous data visualization in relation to their prior beliefs on each topic to assess confirmation bias. On the slider, a value of 0 indicates a neutral stance towards the statement in the prompt. Therefore, the confirmation bias score was calculated using the formula:

$$bias\ score = \pm |evaluation| \tag{4}$$

In this equation, $evaluation$ is the participant's rating of how much the ambiguous data visualization either supports or refutes the statement given in the prompt. The bias score is assigned a positive (+) sign when the evaluation of the visualization aligns with the participant's prior beliefs and a negative (-) sign when it contradicts those beliefs.

Participants. We recruited 1,080 participants from Prolific (M_{age} = 38.79, SD_{age} = 12.96; 531 Male, 537 Female, 5 Non-binary, 2 Transgender, 1 Genderqueer, 1 Genderfluid, 3 Prefer not to disclose). We screened for participants who are based in the United States who are fluent in English with study approval rate greater than 95%. Individuals who participated in preliminary studies were ineligible to participate in the main study. The study plan was reviewed and

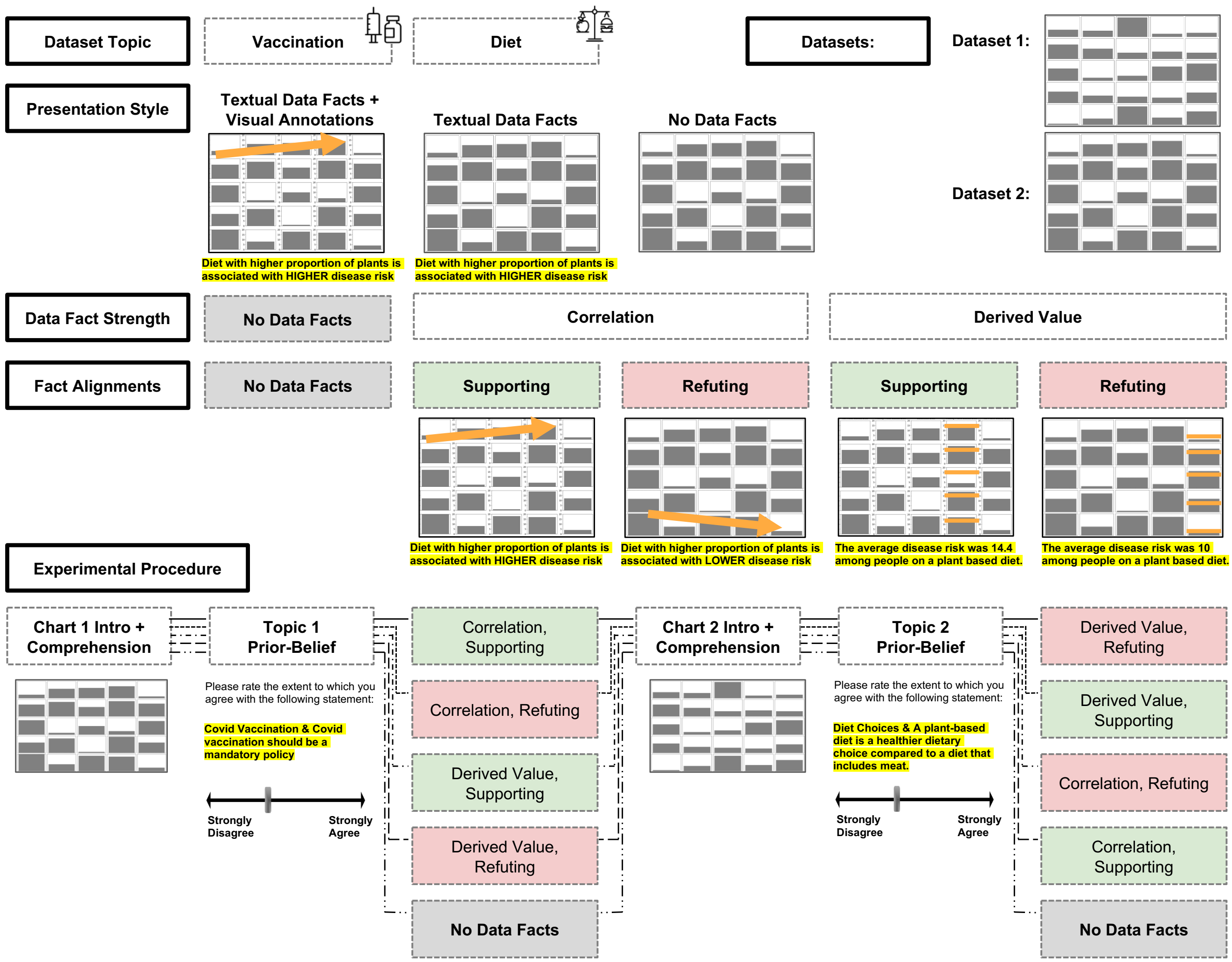

Figure 4: Main Study Conditions and Procedure: [Presentation Style] is a between-subjects condition. Each participant completes 2 trials, either involving 2 data visualizations *with data facts*, where [Dataset, Dataset Topic, Data Facts Strength, Fact Alignment] are within-subject conditions, or 2 visualizations with *No Data Facts*, with [Dataset Topic] as the only within-subject condition. The varying types of strokes connecting the components in the main study procedure diagram represent the different conditions assigned to each participant.

approved by the authors' university ethics board. The sample size was determined by a power analysis conducted using G*Power [22] to ensure a minimum power of 0.75 (indicating a 75% chance of detecting a true effect for each hypothesis if it exists) across all effect sizes from our hypotheses, with an alpha level of 0.05. This analysis was based on a pilot study involving 61 participants under the same settings as the main study.

6 Results

See Table 1 for a summary of results.

Methodology. We use a single mixed-effects linear model to test H1, H2, and H3. For H4 and H5, we separately examine the effects of supporting and refuting data facts using two linear models:

- **Situation SN:** Analyzing data that includes only participants who viewed visualizations with ***Supporting*** Data Facts and *No Data Facts*.
- **Situation RN:** Analyzing data that includes only participants who viewed visualizations with ***Refuting*** Data Facts and *No Data Facts*.

This separation is crucial, as we hypothesize that ***supporting*** data facts increase confirmation bias compared to the no data fact condition (our baseline), raising the score above zero. On the other hand, ***refuting*** data facts decrease confirmation bias compared to our baseline, lowering the score below zero. Analyzing these conditions together could obscure key differences in the distributions due to the effects evening out.

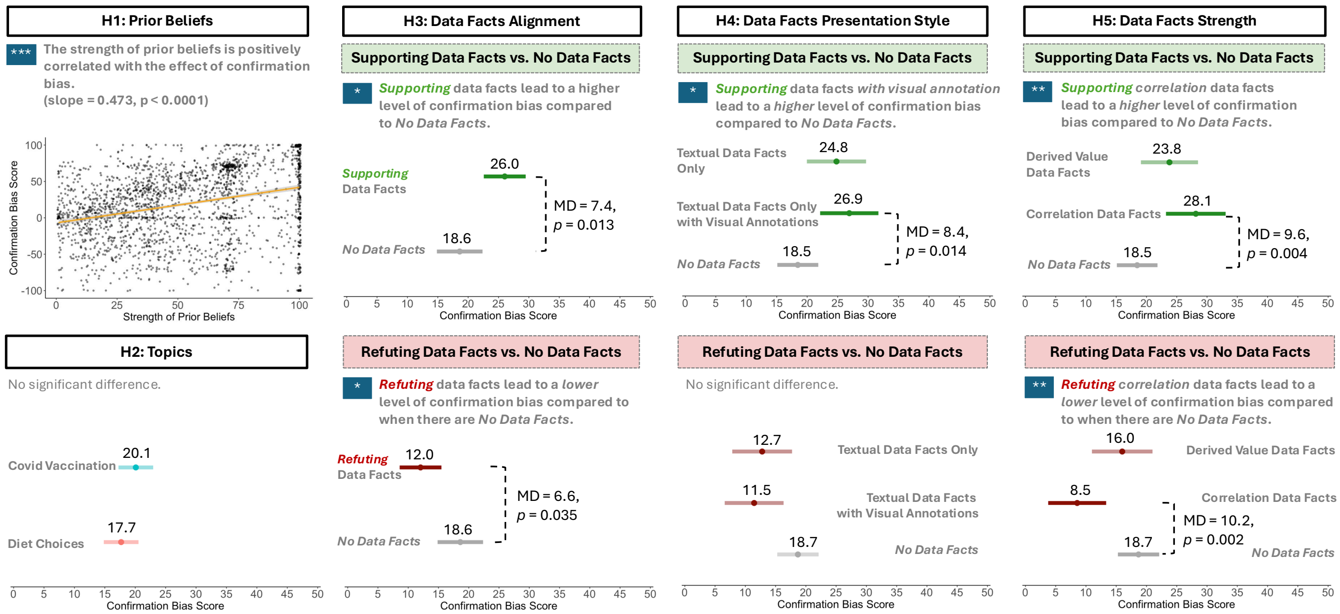

Figure 5: Main Study Hypotheses and Results. The results of H1 are illustrated as a scatterplot showing the relationship between the strength of prior beliefs and confirmation bias. The *estimated marginal means* from each linear model and the 95% confidence intervals of confirmation bias scores across different conditions in H2 - H5 are plotted as confidence interval plots. For H2, colors represent different topics, while for H3, H4, and H5, dark red indicates *refuting* data facts, forest green represents *supporting* data facts, and grey signifies no data facts.

6.1 Analysis of H1, H2 and H3

We created a mixed-effect model to analyze the data. The fixed effects included:

- the strength of initial belief, as calculated from Equation 3 (a continuous integer from 0 to 100),
- the data topic (COVID-19 vaccination or diet choices)
- the data facts alignments (***Supporting***, ***Refuting***, or the reference level: ***No Data Facts***).

Participant ID is the only random effect. The dependent variable is the amount of confirmation bias, as calculated from Equation 4. The results are summarized in Table 3.

H1: Prior Beliefs

As shown in Figure 5, under **H1** *we found that participants' confirmation bias score is significantly positively correlated with the strength of their pre-existing beliefs, regardless of the direction of data facts.* With one unit increase in the strength of the pre-existing belief, the confirmation bias score increases by about 0.473 units. This **supports H1**.

H2: Topic

While we observe a trend that the COVID-19 vaccination topic resulted in higher confirmation bias ($Estimated\ Marginal\ Mean(EMM) = 20.1$) compared to the diet choice topic ($EMM = 17.7$), we observe no significant effect of topic. Thus H2 **is not supported**. One possible explanation for the lack of significance could be that the topical differences are captured by prior beliefs (**H1**). In particular, people hold stronger prior beliefs towards polarizing topics (p < 0.001, 95% CI for the mean difference in initial belief strength between diet and vaccination topics: [-18.27, -13.30]). As a result, when both prior belief strength and topic are included in the model, the topic itself might not significantly impact confirmation bias.

H3: Data Facts Alignment

Compared to the ***No Data Facts*** condition, participants exhibited significantly more confirmation bias when reading ***Supporting Data Facts*** ($EMM\boldsymbol{sup} - EMM\boldsymbol{n} = 7.4$), while they exhibited significantly less confirmation bias when reading ***Refuting Data Facts*** ($EMM\boldsymbol{n} - EMM_{\boldsymbol{ref}} = 6.6$). These results **support H3**.

6.2 Analysis of H4

We conducted the statistical analysis under both situations **SN** and **RN**, as outlined at the beginning of Section 6. Since this is a between-subject comparison, we constructed a linear model with the following predictors:

- the strength of initial belief (a continuous integer from 0 to 100, calculated using Equation 2),
- the topic of the dataset (COVID-19 vaccination or diet choices)
- the data facts presentation style (textual data facts only, textual data facts with visual annotation, or ***No Data Facts***).

The dependent variable was the confirmation bias score. The results are summarized in Table 4.

We found a significant effect of display modes on confirmation bias under both **SN** and **RN**. For the situations where participants

Table 3: Mixed-effect linear modeling results. The fourth column, 'Coefficient P Value,' represents the p-value obtained from the model fitting process for each category under each independent variable. The fifth column, 'Variable P Value,' represents the p-value derived from the Type III ANOVA for each independent variable.

Independent Variable	Coefficient	Std. Error	t-value	Coefficient p-Value	Variable p Value
Initial Belief Strength	0.473	0.034	13.771	< 0.001 ***	< 0.001 ***
Topic	2.414	1.867	1.293	0.196	0.196
Alignment: Refuting	-6.570	2.600	-2.527	0.012 *	Alignment: < 0.001 ***
Alignment: Supporting	7.417	2.600	2.853	0.004 **	

read the ***supporting*** data facts **SN**, post-hoc analysis with Bonferroni correction [70] suggests that the significant effect is driven by the significant difference in confirmation bias levels between the baseline condition (***No Data Facts***) and [Textual Data Facts with Visual Annotations] condition ($p_{\mathbf{SN}} = 0.014$).

For the situations where participants read the ***refuting*** data facts **RN**, post-hoc analysis suggests only a trending significant difference in confirmation bias level between the baseline and the [Textual Data Facts with Visual Annotations] condition ($p_{\mathbf{RN}} = 0.051$).

We observed no significant difference between the baseline and the [Textual Data Facts Only] condition under both **SN** and **RN**. These results **partially support H4**.

6.3 Analysis of H5

We again consider both the situations **SN** and **RN**. We constructed a linear model with the following fixed effects:

- the strength of initial belief (a continuous integer from 0 to 100, calculated using Equation 2)
- the topic of the dataset (COVID-19 vaccination or diet choices)
- the strength of data facts ([Correlation] Data Facts, [Derived Value] Data Facts, or ***No Data Facts***)

The dependent variable was the confirmation bias score. The results are summarized in Table 4.

We found a significant effect of the data fact strength on confirmation bias under both **SN** and **RN**. Post-hoc analysis with Bonferroni correction suggests that for both ***Supporting*** and ***Refuting*** data facts, participants who read the [Correlation] data fact exhibited a significantly **higher** amount of confirmation bias compared to the baseline ($p_{\mathbf{SN}} = 0.004$, $p_{\mathbf{RN}} = 0.002$).

On the other hand, for both ***Supporting*** and ***Refuting*** data facts, participants who read the [Derived Value] data fact did not exhibit a significantly different amount of confirmation bias compared to the baseline. These results **partially support H5**.

7 Discussion

7.1 The Role of Visual Annotation

In our findings on **H4**, we were surprised to observe that ***supporting*** data facts might only increase confirmation bias when accompanied by visual annotations, suggesting that supporting data facts alone may not significantly bias users. This unexpected finding motivates us to reconsider the role of *visual annotations* in visualization design.

The use of annotations and embellishments in visualizations has long been debated within the visualization community. While some researchers advocate for maximizing the data-ink ratio to reduce chart junk that could distract or mislead viewers [60, 61], others argue that annotations can improve the accuracy of information recalled from visualizations [5, 27]. More recent research demonstrated that annotations can significantly influence people to draw conclusions aligned with the annotations, beyond what might naturally capture their attention [6, 57]. These insights together provide a potential explanation for our findings on **H4**: namely, visual annotations could enhance the salience of patterns described by textual data facts and thus shape how individuals interpret a visualization as ***supporting*** or ***refuting*** evidence, which ultimately affects the manifestation of confirmation bias. Therefore, designers should be aware that when visual annotations are combined with textual summaries, there is a strong potential to consolidate viewers' prior beliefs when there is alignment, making individuals more susceptible to confirmation bias.

7.2 (De-)Biased or Echoing?

The opposite effects of ***supporting*** and ***refuting*** data on confirmation bias prompt us to consider potential driving factors. Were people prompted to think more critically when they read ***refuting*** data facts, or did they just echo back the trends described by the data facts? Our results demonstrated that ***refuting*** data facts cause people's beliefs to shift 35% (See Figure 5 H3 Bottom)[2], resulting in an estimated marginal mean of confirmation bias scores of 12.0, significantly greater ($p = 0.035$, 95% CI [8.59, 15.5]) than a baseline of zero (which means no confirmation bias). If people were simply echoing the perspective of the ***refuting*** data facts, we would expect greater belief movement in the opposite direction, with sub-zero values for the confirmation bias score. But this was not the case. In other words, people still exhibited confirmation bias (albeit less) even when viewing ***refuting*** data facts . This finding supports the case that ***refuting*** data facts help mitigate confirmation bias, rather than the case where participants just echo the trends described by the data facts. This observation also corroborates that confirmation bias is a weighting bias rather than an interpretation bias [36], where people assign higher weight to evidence that ***supports*** their prior beliefs than to evidence that ***refutes*** them.

[2] 35% is calculated from the estimated marginal means in Figure 5 H3 Bottom as (18.6-12)/18.6.

Table 4: Fixed-effects linear modeling results. The sixth column, 'Coefficient P Value,' represents the p-value obtained from the model fitting process for each category under each independent variable. The seventh column, 'Variable P Value,' represents the p-value derived from the Type III ANOVA for each independent variable.

Situation	Independent Variable	Coefficient	Std. Error	t-value	Coefficient p	Variable p
SN	Style: Textual Only	6.348	3.011	2.108	0.035 *	Style: 0.008 **
	Style: Textual + Visual	8.451	2.980	2.836	0.005 **	
	Strength: Correlation	9.688	3.039	3.187	0.001 **	Strength: 0.005 **
	Strength: Derived Value	5.333	2.952	1.807	0.071	
RN	Style: Textual Only	-5.905	3.061	-1.929	0.054	Style: 0.0279 *
	Style: Textual + Visual	-7.229	3.029	-2.387	0.017 *	
	Strength: Correlation	-10.134	2.997	-3.382	0.001 **	Strength: 0.003 **
	Strength: Derived Value	-2.694	3.086	-0.873	0.383	

7.3 Exacerbating or Mitigating Confirmation Bias?

Consider a viewer who holds a strongly positive belief about COVID-19 vaccination. If they encounter a data visualization with data facts that ***support*** their existing belief, it could strengthen their conviction by 40% (See Figure 5 H3 Top) However, if they are presented with data facts that ***refute*** their prior belief, it could weaken their conviction, shifting their belief by 35% in the opposite direction (See Figure 5 H3 Bottom)[3]. This suggests that ***refuting*** data facts might help reduce confirmation bias. While the distribution of bias scores with ***refuting*** data facts remains positive (suggesting that some confirmation bias is still present), it is nevertheless lower than in scenarios with ***No Data Facts***, indicating a potential trend toward reducing confirmation bias with ***refuting*** data facts relative to the baseline level of bias. Our findings corroborate previous research in data visualization that highlights the power of text alongside data visualization to shape individuals' interpretations of data [28, 30, 57]. These findings might also indicate that providing alternative interpretations of the same visualization—especially those that challenge prior beliefs—might offer a more balanced perspective of the data and might prevent decisions based solely on ***supporting*** evidence.

Data Facts Consideration Our experimental findings revealed that data facts have potential both to exacerbate or mitigate confirmation bias depending on whether they are framed as ***supporting*** or ***refuting*** the viewer's prior beliefs. Realistically, as the use of large language models proliferates through more application areas, it will be hard to avoid the inevitable combinations of visualizations supported by automatically (or manually) generated textual summaries. Rather, we assert that this is an opportunity to leverage the mitigating potential of data facts in future systems. In particular, if future systems leverage belief elicitation techniques [32] prior to showing users data, systems can intelligently display refuting data facts alongside visualizations, thereby mitigating confirmation bias in interactive settings. Nevertheless, there is a clear risk that malicious actors could misuse these insights to induce a favorable bias. We encourage responsible use of these insights, not as a tool for unfair manipulation, but as a tool to promote balanced thinking and informed decision making.

[3]40% is calculated from the estimated marginal means in Figure 5 H3 Top as (26-18.6)/18.6, and 35% from the means in Figure 5 H3 Bottom as (18.6-12)/18.6.

7.4 Refuting Data Facts and Counterfactuals in Visualizations

Existing work in visualizations have examined other approaches to reduce confirmation bias in visual data analysis. One such approach involves the use of counterfactuals, which refer to the mental constructs that represent *alternatives* to reality [8, 53]. When people think "if only" or "what if," they imagine how the past could have unfolded differently. This process helps reduce bias by encouraging individuals to consider alternative factors that could have led to the same outcome. In an analytical scenario where the data suggests two variables might be causally related, counterfactuals encourage users to question whether the two factors are truly causally related or if a third variable might be influencing the results [66–68]. In another example, highlighting points that have previously been examined (thereby visually distinguishing points that have/have not been previously examined) when analyzing data with visualizations encourages users to focus more on unexplored areas to potentially gain new insights [23, 65]. Further, providing people with ***refuting*** data facts, as we did in our study, can also prompt people to consider alternative perspectives. Thus, counterfactuals, highlighted data values, and ***refuting*** data facts offer users alternative perspectives that challenge their existing beliefs about the data. These insights underscore the potential de-biasing power of incorporating multiple data perspectives within a visualization. By empowering analysts to consider a wider range of alternatives for a balanced analysis, these techniques can combat over-reliance on existing beliefs and reduce confirmation bias.

7.5 Limitations

One limitation of our study is the narrow set of topics tested. We selected only one polarizing topic and one neutral topic, with their levels of polarization evaluated by a small group of 40 participants. However, there are other topics where people overwhelmingly share the same stance, such as the general affection people have for pets [3]. Future research should test a wider range of topics to further understand the generalizability of our findings.

Another limitation is the use of synthetic datasets. While these datasets allowed us maximum control over our experiment, they may not fully capture real-world scenarios relevant to the topics studied. For example, a participant who is particularly savvy in the topic may suspect the authenticity of the dataset and evaluate associated evidence in unpredictable ways. Additionally, in the present study, we operationalized confirmation bias as interpreting ambiguous information in a way that aligns with one's prior belief, and thus, designed visualizations to be intentionally ambiguous. In the real world, not all visualziations would be designed to be ambiguous; some are explicitly crafted to highlight strong data patterns and convey a particular narrative. Confirmation bias would undoubtedly happen with these visualizations as well. Existing work by Xiong et al [72] have demonstrated that in these cases, visual representation formats such as showing the data via a table or a bar, would change the strength of confirmation bias. Therefore, we encourage future work to also investigate the effect of data facts and visual annotations on confirmation bias in these scenarios where the visualization stimuli are not ambiguous.

Furthermore, the real-world factors related to the statements in Table 2 are likely more complex than what is depicted in the data. For instance, in the COVID-19 Vaccination topic, participants might also consider factors such as mortality rates across different age groups to determine if the vaccination should be mandatory, while in the Diet Choices topic, other health metrics beyond the risk of chronic disease—such as Body Mass Index (BMI), nutrition levels, and emotional well-being—are also indicators of diet health when making comparisons. The limited factors provided in our data might introduce distortions in users' evaluations that are, to some extent, inevitable in controlled studies. Future work can incorporate more realistic datasets to generalize findings in ecologically valid settings that better reflect the complexity of real-world decision-making processes.

Finally, we also determined the strength of the data facts solely based on modulating which visual patterns they described (*PS 3*, Section 4.3). Further research could explore how factors such as the phrasing of the data facts, source credibility, and trustworthiness influence their perceived strength.

8 Conclusion

Across a series of experiments, we explored the impact of presenting textual data facts alongside visualizations and uncovered key insights about confirmation bias. Participants with stronger-held prior beliefs were more likely to interpret ambiguous visualizations as supporting their prior beliefs. When visualizations were accompanied by textual data facts, confirmation bias was *exacerbated* (when visualizations were paired with ***supporting*** data facts) and *mitigated* (when paired with ***refuting*** data facts). Confirmation bias was further exacerbated when data facts focused on *correlations* compared to *average values*. Our results also suggest that visually annotating ***supporting*** data facts further amplified these effects. These findings together indicate that both the design of the accompanying data facts and an individual's prior beliefs play crucial roles in shaping how visual information is interpreted. Our research highlights the complex interplay between visual and textual evidence in shaping data interpretation. Particularly with recent advances in AI-generated content, we hope these results will stimulate further discussion and serve as a catalyst for a research agenda within the community to assess the potential benefits and drawbacks of combining visualizations with data facts.

Acknowledgments

This work was supported by the National Science Foundation under grants IIS-2311574, IIS-2237585, and IIS-2311575.

References

[1] [n. d.]. Google Charts Introduction. https://developers.google.com/chart/. Accessed: 2024-09-10.

[2] Robert Amar, James Eagan, and John Stasko. 2005. Low-level components of analytic activity in information visualization. In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.* IEEE, 111–117.

[3] John Archer. 1997. Why do people love their pets? *Evolution and Human behavior* 18, 4 (1997), 237–259.

[4] Calvin S. Bao, Siyao Li, Sarah G Flores, Michael Correll, and Leilani Battle. 2022. Recommendations for Visualization Recommendations: Exploring Preferences and Priorities in Public Health. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 411, 17 pages. doi:10.1145/3491102.3501891

[5] Scott Bateman, Regan L Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher Brooks. 2010. Useful junk? The effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2573–2582.

[6] Cindy Xiong Bearfield, Lisanne van Weelden, Adam Waytz, and Steve Franconeri. 2024. Same Data, Diverging Perspectives: the Power of Visualizations to Elicit Competing Interpretations. In *2024 IEEE Pacific Visualization*. IEEE.

[7] Laura A Brannon, Michael J Tagler, and Alice H Eagly. 2007. The moderating role of attitude strength in selective exposure to information. *Journal of Experimental Social Psychology* 43, 4 (2007), 611–617.

[8] Ruth MJ Byrne. 2016. Counterfactual thought. *Annual review of psychology* 67, 1 (2016), 135–157.

[9] Steve D Charman, Melissa Kavetski, and Dana Hirn Mueller. 2017. Cognitive bias in the legal system: Police officers evaluate ambiguous evidence in a belief-consistent manner. *Journal of Applied Research in Memory and Cognition* 6, 2 (2017), 193–202.

[10] Muhammad Faisal Cheema, Stefan Jänicke, and Gerik Scheuermann. 2016. AnnotateVis: Combining traditional close reading with visual text analysis. In *Workshop on Visualization for the Digital Humanities, IEEE VIS*.

[11] Tee Chuanromanee and Ronald Metoyer. 2022. A Crowdsourced Study of Visual Strategies for Mitigating Confirmation Bias. In *2022 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. IEEE, 1–6.

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

[13] Michael Correll. 2019. Ethical dimensions of visualization research. In *Proc. of the 2019 CHI Conf. on Human Factors in Comput. Sys.* 1–13.

[14] Ariel Deardorff. 2016. Tableau (version. 9.1). *Journal of the Medical Library Association* 104, 2 (2016), 182–183.

[15] Ritson Delpish, Steven Jiang, Lauren Davis, and Kehinde Odubela. 2019. A visual analytics approach to combat confirmation bias for a local food bank. In *Advances in Human Error, Reliability, Resilience, and Performance: Proceedings of the AHFE 2018 International Conference on Human Error, Reliability, Resilience, and Performance, July 21-25, 2018, Loews Sapphire Falls Resort at Universal Studios, Orlando, Florida, USA 9*. Springer, 13–23.

[16] David L Dickinson. 2022. Political ideology, mood response, and the confirmation bias. (2022).

[17] Evanthia Dimara, Steven Franconeri, Catherine Plaisant, Anastasia Bezerianos, and Pierre Dragicevic. 2018. A task-based taxonomy of cognitive biases for information visualization. *IEEE transactions on visualization and computer graphics* 26, 2 (2018), 1413–1432.

[18] Rui Ding, Shi Han, Yong Xu, Haidong Zhang, and Dongmei Zhang. 2019. QuickInsights: Quick and Automatic Discovery of Insights from Multi-Dimensional Data. In *Proceedings of the 2019 International Conference on Management of Data* (Amsterdam, Netherlands) *(SIGMOD '19)*. Association for Computing Machinery, New York, NY, USA, 317–332. doi:10.1145/3299869.3314037

[19] Hamza Elhamdadi, Adam Stefkovics, Johanna Beyer, Eric Moerth, Hanspeter Pfister, Cindy Xiong Bearfield, and Carolina Nobre. 2023. Vistrust: a Multidimensional Framework and Empirical Study of Trust in Data Visualizations. *IEEE Transactions on Visualization and Computer Graphics* (2023).

[20] Madison A Elliott, Christine Nothelfer, Cindy Xiong, and Danielle Albers Szafir. 2020. A design space of vision science methods for visualization research. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2020), 1117–1127.

[21] Jonathan St BT Evans. 1989. *Bias in human reasoning: Causes and consequences.* Lawrence Erlbaum Associates, Inc.

[22] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods* 39, 2 (2007), 175–191.

[23] Mi Feng, Cheng Deng, Evan M Peck, and Lane Harrison. 2016. Hindsight: Encouraging exploration through direct encoding of personal interaction history. *IEEE transactions on visualization and computer graphics* 23, 1 (2016), 351–360.

[24] Alberto Ferrari and Marco Russo. 2016. *Introducing Microsoft Power BI.* Microsoft Press.

[25] Marti A Hearst. 2023. Show It or Tell It? Text, Visualization, and Their Combination. *Commun. ACM* 66, 10 (2023), 68–75.
[26] Andreas Hergovich, Reinhard Schott, and Christoph Burger. 2010. Biased evaluation of abstracts depending on topic and conclusion: Further evidence of a confirmation bias within scientific psychology. *Current Psychology* 29 (2010), 188–209.
[27] Jessica Hullman, Eytan Adar, and Priti Shah. 2011. Benefitting infovis with visual difficulties. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2213–2222.
[28] Dae Hyun Kim, Vidya Setlur, and Maneesh Agrawala. 2021. Towards Understanding How Readers Integrate Charts and Captions: A Case Study with Line Charts. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11.
[29] Joshua Klayman. 1995. Varieties of confirmation bias. *Psychology of learning and motivation* 32 (1995), 385–418.
[30] Ha-Kyung Kong, Zhicheng Liu, and Karrie Karahalios. 2018. Frames and slants in titles of visualizations on controversial topics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.
[31] Ha-Kyung Kong, Zhicheng Liu, and Karrie Karahalios. 2019. Trust and recall of information across varying degrees of title-visualization misalignment. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.
[32] Ratanond Koonchanok, Gauri Yatindra Tawde, Gokul Ragunandhan Narayanasamy, Shalmali Walimbe, and Khairi Reda. 2023. Visual belief elicitation reduces the incidence of false discovery. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.
[33] Chufan Lai, Zhixian Lin, Ruike Jiang, Yun Han, Can Liu, and Xiaoru Yuan. 2020. Automatic annotation synchronizing with textual description for visualization. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
[34] Shahid Latif, Zheng Zhou, Yoon Kim, Fabian Beck, and Nam Wook Kim. 2022. Kori: Interactive Synthesis of Text and Charts in Data Documents. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2022), 184–194. doi:10.1109/TVCG.2021.3114802
[35] Po-Ming Law, Alex Endert, and John Stasko. 2020. Characterizing automated data insights. In *2020 IEEE Visualization Conference (VIS)*. IEEE, 171–175.
[36] Paul Edward Lehner, Leonard Adelman, Brant A Cheikes, and Mark J Brown. 2008. Confirmation bias in complex analyses. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 38, 3 (2008), 584–592.
[37] Quan Li, Huanbin Lin, Chun Feng Tang, Xiguang Wei, Zhenhui Peng, Xiaojuan Ma, and Tianjian Chen. 2021. Exploring the" Double-Edged Sword" Effect of Auto-Insight Recommendation in Exploratory Data Analysis.. In *IUI Workshops*.
[38] Rich Ling. 2020. Confirmation bias in the era of mobile news consumption: the social and psychological dimensions. *Digital Journalism* 8, 5 (2020), 596–604.
[39] Can Liu, Liwenhan Xie, Yun Han, Datong Wei, and Xiaoru Yuan. 2020. AutoCaption: An approach to generate natural language description from visualization automatically. In *2020 IEEE Pacific visualization symposium (PacificVis)*. IEEE, 191–195.
[40] Douglas Markant, Milad Rogha, Alireza Karduni, Ryan Wesslen, and Wenwen Dou. 2023. When do data visualizations persuade? The impact of prior attitudes on learning about correlations from scatterplot visualizations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.
[41] Isaac D Mehlhaff. 2022. A Group-Based Approach to Measuring Polarization. *American Political Science Review* (2022), 1–9.
[42] Atilla Alpay Nalcaci, Dilara Girgin, Semih Balki, Fatih Talay, Hasan Alp Boz, and Selim Balcisoy. 2019. Detection of Confirmation and Distinction Biases in Visual Analytics Systems.. In *TrustVis@ EuroVis*. 13–17.
[43] Raymond S Nickerson. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology* 2, 2 (1998), 175–220.
[44] Alexander Nussbaumer, Katrien Verbert, Eva-Catherine Hillemann, Michael A Bedek, and Dietrich Albert. 2016. A framework for cognitive bias detection and feedback in a visual analytics environment. In *2016 European Intelligence and Security Informatics Conference (EISIC)*. IEEE, 148–151.
[45] Jason Obeid and Enamul Hoque. 2020. Chart-to-text: Generating natural language descriptions for charts by adapting the transformer model. *arXiv preprint arXiv:2010.09142* (2020).
[46] Stefan Palan and Christian Schitter. 2018. Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance* 17 (2018), 22–27.
[47] Anshul Vikram Pandey, Anjali Manivannan, Oded Nov, Margaret Satterthwaite, and Enrico Bertini. 2014. The persuasive power of data visualization. *IEEE transactions on visualization and computer graphics* 20, 12 (2014), 2211–2220.
[48] Saugat Pandey, Oen G McKinley, R Jordan Crouser, and Alvitta Ottley. 2023. Do You Trust What You See? Toward A Multidimensional Measure of Trust in Visualization. *arXiv preprint arXiv:2308.04727* (2023).
[49] Marianne Procopio, Ab Mosca, Carlos Scheidegger, Eugene Wu, and Remco Chang. 2021. Impact of cognitive biases on progressive visualization. *IEEE Transactions on Visualization and Computer Graphics* 28, 9 (2021), 3093–3112.
[50] Deepthi Raghunandan, Zhe Cui, Kartik Krishnan, Segen Tirfe, Shenzhi Shi, Tejaswi Darshan Shrestha, Leilani Battle, and Niklas Elmqvist. 2022. Lodestar: Supporting independent learning and rapid experimentation through data-driven analysis recommendations. *arXiv preprint arXiv:2204.07876* (2022).
[51] Donghao Ren, Matthew Brehmer, Bongshin Lee, Tobias Höllerer, and Eun Kyoung Choe. 2017. Chartaccent: Annotation for data-driven storytelling. In *2017 IEEE Pacific Visualization Symposium (PacificVis)*. Ieee, 230–239.
[52] María Teresa Rodríguez, Sérgio Nunes, and Tiago Devezas. 2015. Telling stories with data visualization. In *Proceedings of the 2015 workshop on narrative & hypertext*. 7–11.
[53] Neal J Roese. 1997. Counterfactual thinking. *Psychological bulletin* 121, 1 (1997), 133.
[54] Walter R Schumm. 2021. Confirmation bias and methodology in social science: an editorial. 285–293 pages.
[55] Edward Segel and Jeffrey Heer. 2010. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics* 16, 6 (2010), 1139–1148.
[56] Arjun Srinivasan, Steven M Drucker, Alex Endert, and John Stasko. 2018. Augmenting visualizations with interactive data facts to facilitate interpretation and communication. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 672–681.
[57] Chase Stokes, Cindy Xiong Bearfield, and Marti A Hearst. 2023. The Role of Text in Visualizations: How Annotations Shape Perceptions of Bias and Influence Predictions. *IEEE Transactions on Visualization and Computer Graphics* (2023).
[58] Chase Stokes, Vidya Setlur, Bridget Cogley, Arvind Satyanarayan, and Marti A Hearst. 2022. Striking a balance: Reader takeaways and preferences when integrating text and charts. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2022), 1233–1243.
[59] Charles S Taber and Milton Lodge. 2006. Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science* 50, 3 (2006), 755–769.
[60] Edward R Tufte. 1991. Envisioning information. *Optometry and Vision Science* 68, 4 (1991), 322–324.
[61] Edward R Tufte and Peter R Graves-Morris. 1983. *The visual display of quantitative information*. Vol. 2. Graphics press Cheshire, CT.
[62] Dáša Vedejová and Vladimíra Čavojová. 2022. Confirmation bias in information search, interpretation, and memory recall: Evidence from reasoning about four controversial topics. *Thinking & Reasoning* 28, 1 (2022), 1–28.
[63] Emily Wall, Leslie Blaha, Celeste Paul, and Alex Endert. 2019. A formative study of interactive bias metrics in visual analytics using anchoring bias. In *Human-Computer Interaction–INTERACT 2019: 17th IFIP TC 13 International Conference, Paphos, Cyprus, September 2–6, 2019, Proceedings, Part II 17*. Springer, 555–575.
[64] Emily Wall, Leslie M Blaha, Lyndsey Franklin, and Alex Endert. 2017. Warning, bias may occur: A proposed approach to detecting cognitive bias in interactive visual analytics. In *2017 ieee conference on visual analytics science and technology (vast)*. IEEE, 104–115.
[65] Emily Wall, Arpit Narechania, Adam Coscia, Jamal Paden, and Alex Endert. 2021. Left, Right, and Gender: Exploring Interaction Traces to Mitigate Human Biases. *IEEE transactions on visualization and computer graphics* (2021).
[66] Arran Zeyu Wang, David Borland, and David Gotz. 2024. An empirical study of counterfactual visualization to support visual causal inference. *Information Visualization* 23, 2 (2024), 197–214.
[67] Arran Zeyu Wang, David Borland, and David Gotz. 2024. A framework to improve causal inferences from visualizations using counterfactual operators. *Information Visualization* (2024), 14738716241265120.
[68] Arran Zeyu Wang, David Borland, and David Gotz. To Appear in 2025. Beyond Correlation: Incorporating Counterfactual Guidance to Better Support Exploratory Visual Analysis. *IEEE Transactions on Visualization and Computer Graphics* 31, 1 (To Appear in 2025).
[69] Yun Wang, Zhida Sun, Haidong Zhang, Weiwei Cui, Ke Xu, Xiaojuan Ma, and Dongmei Zhang. 2019. Datashot: Automatic generation of fact sheets from tabular data. *IEEE transactions on visualization and computer graphics* 26, 1 (2019), 895–905.
[70] Eric W Weisstein. 2004. Bonferroni correction. *https://mathworld. wolfram. com/* (2004).
[71] Axel Westerwick, Benjamin K Johnson, and Silvia Knobloch-Westerwick. 2017. Confirmation biases in selective exposure to political online information: Source bias vs. content bias. *Communication Monographs* 84, 3 (2017), 343–364.
[72] Cindy Xiong, Elsie Lee-Robbins, Icy Zhang, Aimen Gaba, and Steven Franconeri. 2022. Reasoning affordances with tables and bar charts. *IEEE transactions on visualization and computer graphics* (2022).
[73] Cindy Xiong, Chase Stokes, Yea-Seul Kim, and Steve Franconeri. 2022. Seeing What You Believe or Believing What You See? Belief Biases Correlation Estimation. *IEEE Transactions on Visualization and Computer Graphics* (2022). to appear.
[74] Cindy Xiong, Lisanne Van Weelden, and Steven Franconeri. 2019. The curse of knowledge in visual data communication. *IEEE transactions on visualization and computer graphics* 26, 10 (2019), 3051–3062.

[75] Itai Yanai and Martin J Lercher. 2021. Does the data do the talking? Confirmation bias in the interpretation of data visualizations. (2021).