Causal Priors and Their Influence on Judgements of Causality in Visualized Data

DArran Zeyu Wang , David Borland , Tabitha C. Peck , Wenyuan Wang , David Gotz

Read & presented by Martine Jansen, for the Reading Group "Statistics Communication and (in)numeracy"

Article

- **Title** - Causal Priors and Their Influence on Judgements of Causality in Visualized Data
- **Authors** - Arran Zeyu Wang , David Borland , Tabitha C. Peck , Wenyuan Wang , David Gotz
- **Published** - IEEE Transactions on Visualization and Computer Graphics, Jan. 2025, pp. 765-775, vol. 31
- **DOI** - https://doi.ieeecomputersociety.org/10.1109/TVCG.2024.3456381 🔒
- **Version I read** - https://vaclab.unc.edu/publication/tvcg_2025_wang/tvcg_2025_wang.pdf 🔓
- **Data, code etc** - Data & R Code is given
  - See https://osf.io/dfkv4/overview?view_only=f84ffbc28cdf45e5a3d68f2f1e9c8427 ☺
  - R code is 18 lines of code, including 8 comment lines 😮
- 1 of 5 articles funded with total NSF award of $1,215,999.00, https://www.nsf.gov/awardsearch/show-award?AWD_ID=2211845

Main idea

“Effectively communicating causal relationships, or the lack thereof, remains challenging for data visualizations [5]. Understanding how humans interpret causal relationships from visualized data is therefore crucial for effective visualization design”

From the Abstract

- investigate factors contributing to perceiving of causal relationships in visualizations
- people make assumptions about causal relationships between pairs of concepts even without seeing any visualized data
- results suggest that these assumptions constitute causal priors
- in combination with visualized association, these impact how data visualizations are interpreted
- those priors can also impact users' confidence in their causal assessments
- our results align with prior work, indicating that chart type may also affect causal inference

Background

read for http://sannejwwillems.nl/communication-reading-group/

Human Perception in Visual Causal Inference

- Causal inference emerging topic in visualization research
- Lot of viz techniques designed to support causal inference
- Visual causal inferences
  - made by a user based on their perception and interpretation of data seen during an analytical task
  - closely related to but distinct from statistical inference approaches
  - made even when viewing basic statistical charts that may focus on correlation or other non-causal statistical measures

read for http://sannejwwillems.nl/communication-reading-group/

Human Cognition and Underlying Causal Priors

- different verbs imply different types and degrees of causal relationships
- people tend to implicitly associate causal relationships with short phrases or individual words
- causal prior: the idea that a causal relation can be implied by the use of words in and of themselves, even before those words are used within a visualization

“In this way, causal priors are similar to other factors describing prior experience or knowledge that have been shown within the visualization literature to influence how people interpret charts”

Study 1: Examining Causal Priors for Concept Pairs

read for http://sannejwwillems.nl/communication-reading-group/

56 Concept pairs

- 50 pairs of variable names from widely used open-source datasets
- 5 pairs from Spurious Correlations
- 1 with plain Variable A and Variable B

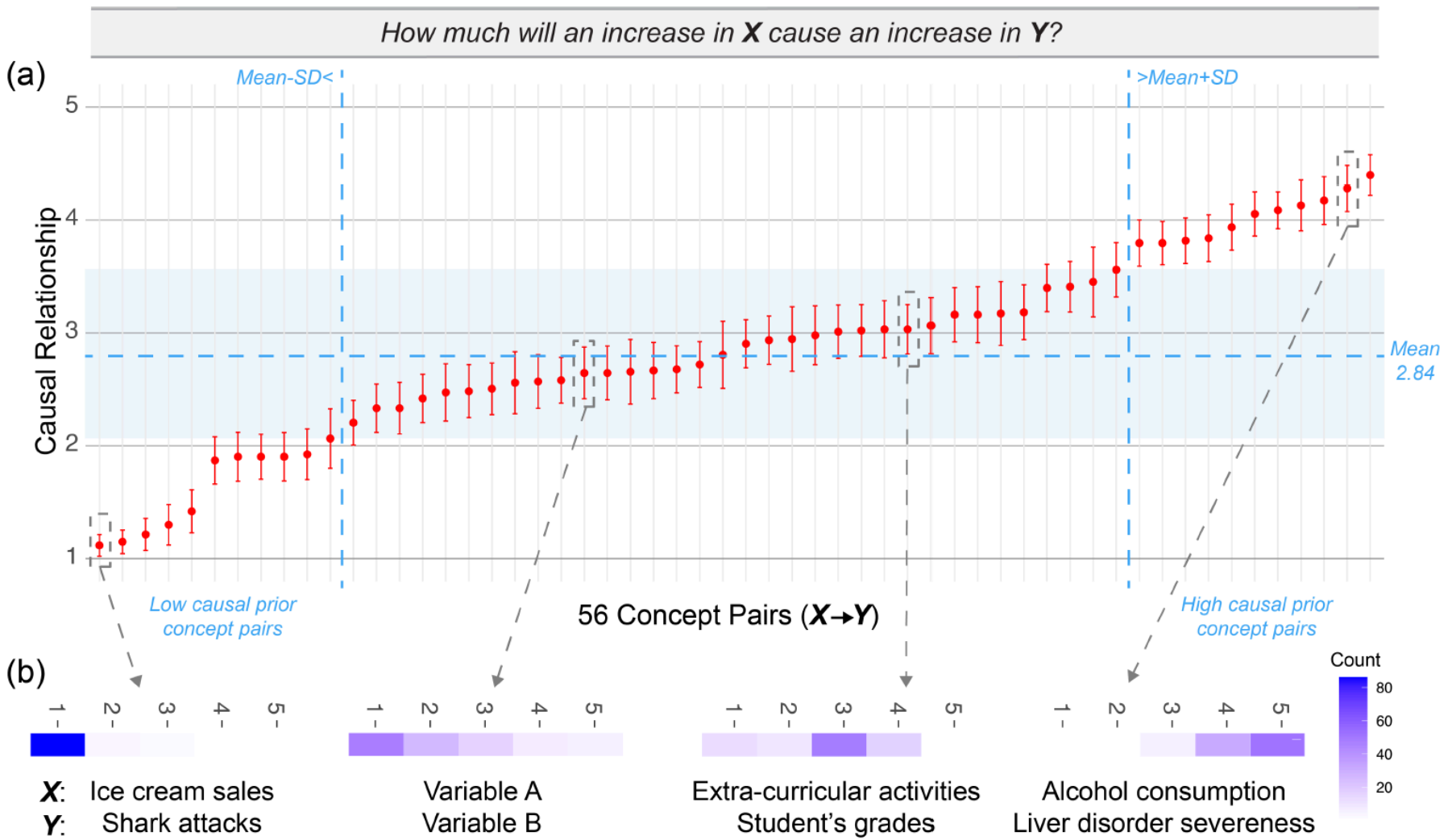
How much will an increase in *age* cause higher *grades*
How much will a person's higher *alcohol consumption* cause more severe *liver disorder*
How much will different *car color* cause different *car prices*
How much will higher *ice cream sales* cause more *shark attacks*
How much will higher *Variable A* cause a higher *Variable B*

92 participants rated

- Amazon MTurk
- asked to rate the strength of the causal relationship
  - on a 5-point Likert scale
  - from 1 (no causal relationship) to 5 (high causal relationship)
- asked to assess their confidence in each rating
  - on a 5-point Likert scale
  - from 1 (low confidence) to 5 (high confidence)

Result: differences between pairs

So there are causal priors



read for http://sannejwwillems.nl/communication-reading-group/

Study 2: Effect of Viz on Causal Perception

read for http://sannejwwillems.nl/communication-reading-group/

56 Items by 229 participants

- 3 categories concept pairs:
  - 22 time series (including the spurious and the A/B)- line charts
  - 17 categorical - bar charts
  - 17 continuous - scatterplots
- 1 of 5 (random) levels of association, with some noise

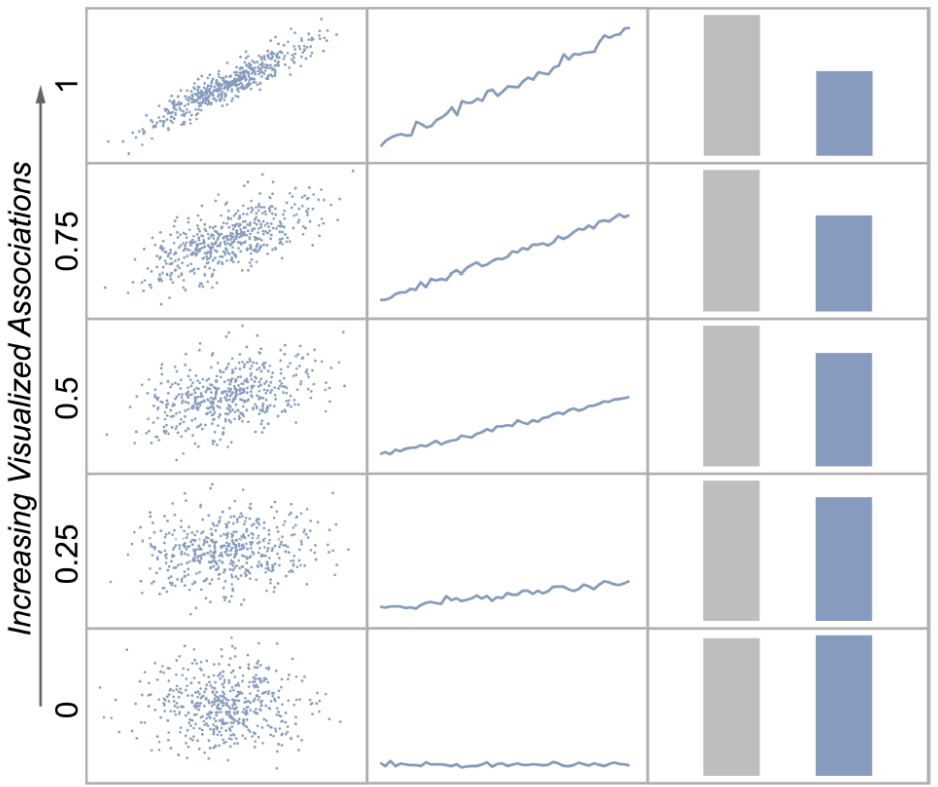
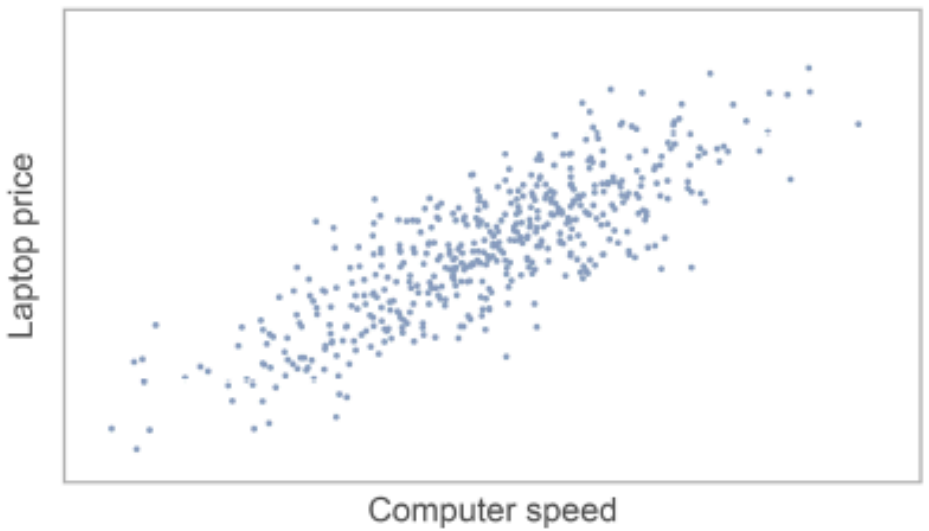


Fig. 3: This figure shows illustrations of the three chart types and five visualized association levels employed in Study 2. Left to right: scatterplots, line charts, and bar charts. Bottom to top: low association (around 0) to high association (around 1).

read for http://sannejwwillems.nl/communication-reading-group/

Example

This chart shows data for **computer speed** and **laptop price**.
Please examine the chart and answer the questions below.



How much will an increase in **computer speed** cause an increase in **laptop price**?

○ 1 (not at all) ○ 2 ○ 3 ○ 4 ○ 5 (entirely)

Please rate your confidence.

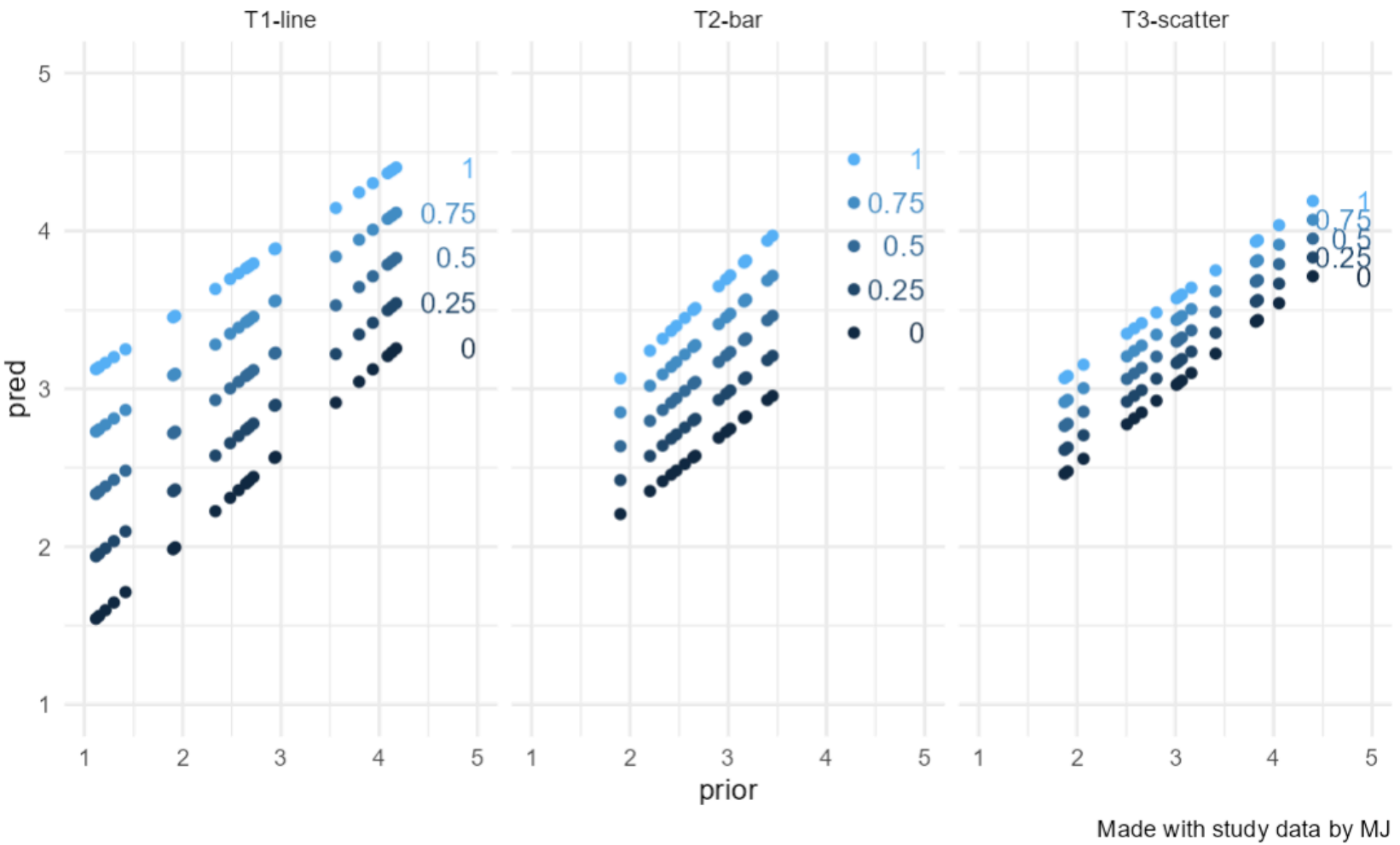
○ 1 (very low) ○ 2 ○ 3 ○ 4 ○ 5 (very high)

Alas no examples of barchart nor linechart

read for http://sannejwwillems.nl/communication-reading-group/

Results

- glm(Score ~ Causal prior ∗ Visual Association ∗ Chart type)



read for http://sannejwwillems.nl/communication-reading-group/

Results

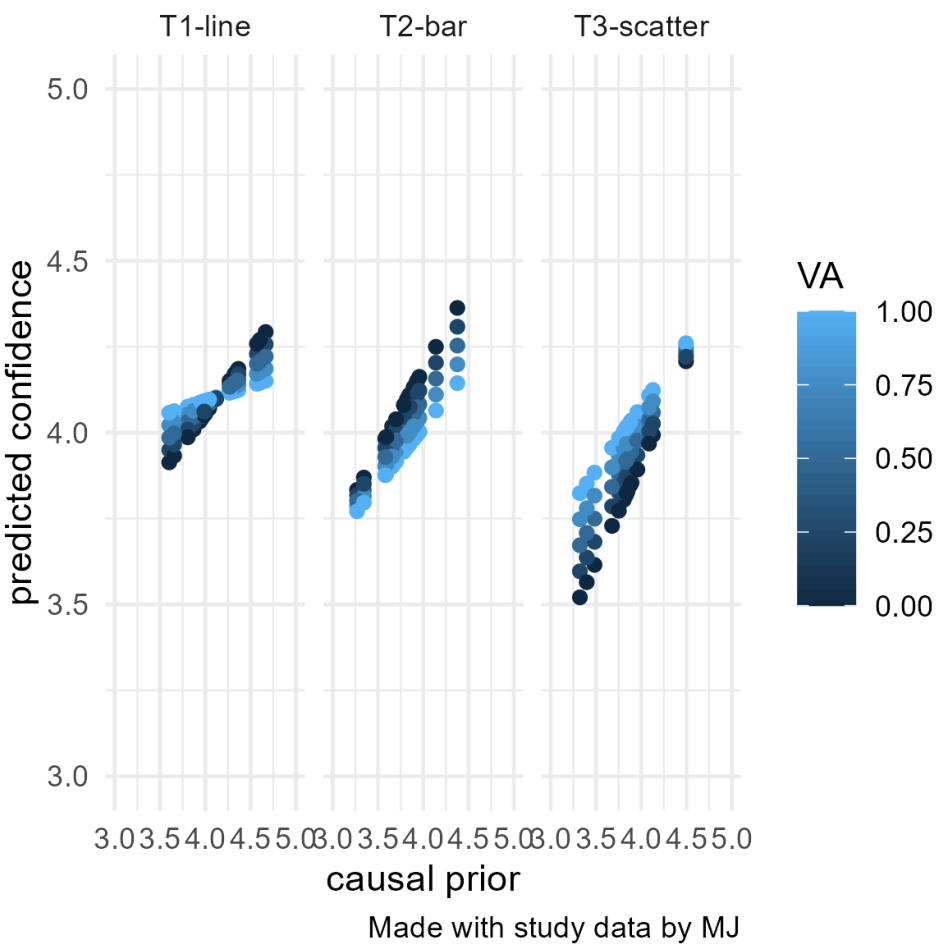
- higher visualized association is associated with higher ratings for strengths of perceived causal relationship
- As causal prior increases, influence of visualized association decreases ?
- visualized association had a greater influence on the perceived causal relationship for line and bar plots than for scatterplots
  - may not be due only to chart type, since chart type ~ data type ☺

Confidence

glm(Score~Prior∗VA∗CT, data = data_s2_confidence)

Similar to the above analysis, we found the generalized linear regression on *confidence* was statistically significant, ($F(11, 14404) = 24.51$, $p < .0001$, $R^2 = .02$), with a significant interaction effect of *causal prior* and *visualized association* ($\beta = -0.27$, $p = .0002$). This suggests a larger disagreement between *causal prior* and *visualized association* may lead to lower subjective *confidence* and vice versa. Please see the OSF supplements for a more comprehensive reporting of these results.

- 😮
- conf score separate from causal score ☹️
- different causal priors than before? 😮



read for http://sannejwwillems.nl/communication-reading-group/

ADDITIONAL RESULTS

Figure 1 shows our overall results from Study 1, for both causal relationships and confidence. This is a merged version of Figure 1 (a) and Figure 6 (a) in the main paper.

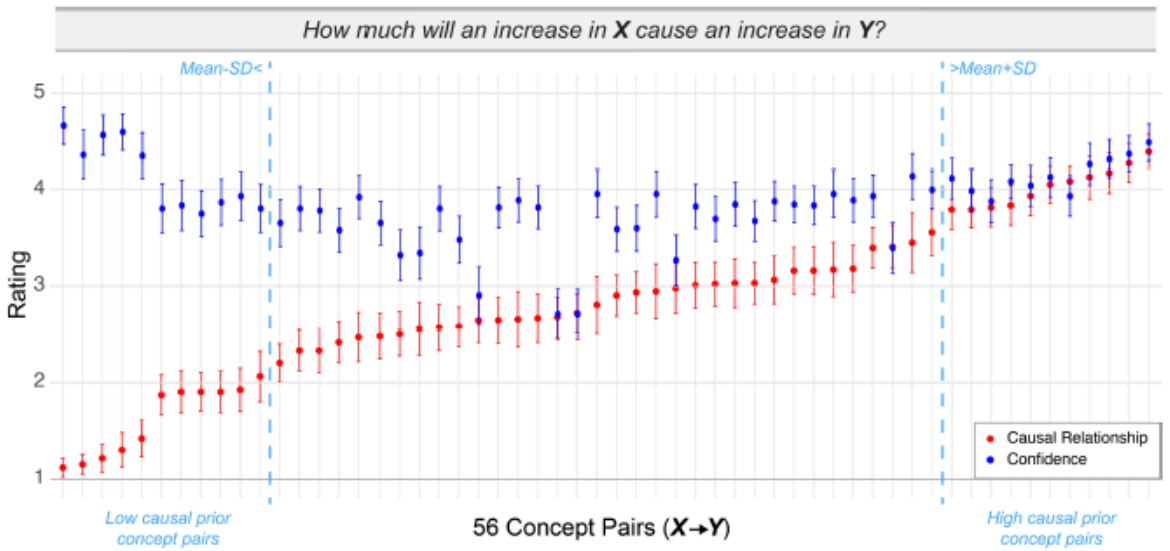


Fig. 1: This figure shows the results of user-reported ratings for both causal relationships (red) and confidence (blue) from Study 1. Same as the Figure 1 in the main paper, the concept pairs are ordered with increasing causal prior and the cyan lines indicate mean ± SD cyan lines.

Table 1 shows our regression analysis results of users' confidence from Study 2. The results are discussed in Section 5.4 of the main paper and the computations are available through the analysis code in *Study.zip*.

Table 1: This figure shows the generalized linear regression results of user-reported ratings of confidence from Study 2.

	β	Std. Error	t-value	p-value
(Intercept)	2.62	0.17	14.56	**< .0001**
Causal Prior	0.35	0.04	7.86	**< .0001**
Visualized Association (VA)	1.11	0.29	3.83	**.0001**
(Chart Type) Bar	-0.35	0.44	-0.81	.41
(Chart Type) Scatterplot	-1.05	0.43	-2.43	**.01**
Causal Prior * VA	-0.26	0.07	-3.66	**.0002**
Bar * Causal Prior	0.12	0.11	1.05	.29
Scatterplot * Causal Prior	0.22	0.11	2.03	**.04**
Bar * VA	-0.71	0.72	-0.98	.32
Scatterplot * VA	-0.10	0.70	-0.14	.88
Bar * Causal Prior * VA	0.12	0.18	0.67	.49
Scatterplot * Causal Prior * VA	0.05	0.18	0.30	.76

Discussion

- “human cognition of causal inference is impacted by both preexisting causal priors and the displayed visualizations”
- “The first major finding indicates that people have preexisting causal priors”
- “our results demonstrate that the correlations broadly represented in visualizations significantly impact human causal inference”
- “We found that causal priors have a significant impact on the cognition of causal inference from visualizations”
- “our results suggest that chart type may impact causal inference” (they mention limitations)

Heuristic implications for design

- Causal priors:
  “future design and research should also consider the causal priors of the users when presenting causal information, and provide ways to elicit, validate, and update them”
- Disagreement impact:
  “Disagreement between a user’s causal prior and the visualized association can lead to cognitive dissonance on overestimation or underestimation of causal relationships. Since the visualized association can be pre-calculated by visual analytics systems, we would recommend adding functionalities of disagreement checks using our suggested model in Section 6.3, if the causal prior has been collected in advance”
- Visualization types:
  “Developers should be aware the overall cognitive process of causal inference that chart types may evoke should be an important consideration of effective visualization design. Future design efforts should focus on identifying which types of visual representations are most conducive to accurate causal inference.”