Title
Processing Scatterplots: Impact of Outliers on Correlational and Causal Inferences

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Author
Boduroglu, Aysecan

Publication Date
2023

Peer reviewed

Processing Scatterplots: Impact of Outliers on Correlational and Causal Inferences

Aysecan Boduroglu
Department of Psychology, Koç University

Abstract

Scatterplot research has identified factors that impact people's perception of correlation magnitudes, yet much less is known about how people reason about data represented in scatterplots. We investigated how people make correlational and causal inferences based on scatterplots with and without outliers. In Experiment 1 and 2, participants viewed scatterplots matched in overall correlational magnitude depicted, but half had an outlier. In Experiment 3, the scatterplots in the two conditions were matched in the correlation magnitude depicted by all the dots excluding the outlier. For each scatterplot, participants stated their endorsement for correlational (X and Y change together) and causal statements (X changes Y). Only when outliers further strengthened an already moderate to strong relationship, people endorsed related correlational statements more and showed a stronger causality bias. Altogether we demonstrate that the impact of outliers in scatterplots on visual reasoning depends on the strength of the relationship depicted.

Keywords: scatterplot; outlier; causality bias

Introduction

In both scientific and everyday communication of data, graphs are utilized more than other data visualizations. It is well known that graphs can be effective in communicating various risks (Fagerlin et al., 2011; Lipkus & Hollands, 1999), financial information (Beattie & Jones, 2008; Merkl-Davies & Brennan, 2007) and evolving global trends, as we clearly saw in the recent pandemic (Charterjee et al., 2021; Fansher et al., 2022a; Romano et al., 2020; Zacks & Franconeri, 2021). However, it is also known that graphs can be misleading either because people do not have the levels of graph literacy to ensure their effective processing (Galesic & Garcia-Retamero, 2011; Lee et al.,2017), or because graph formats lead to various interpretive biases (Fansher et al., 2022b; Shah & Hoeffner, 2002; Tumen & Boduroglu, 2022). There are also some controversial findings as to whether graphs increase or hinder scientific reasoning (e.g. Tal, 2015; Tal & Wansick, 2016, but also see Dragisevic & Jansen 2015; Fansher et al., 2022c).

One major class of scientific reasoning errors is known as the causality bias, the tendency to erroneously make inferences of causality from correlational data (Shah et al., 2017). It has been shown that people often conclude that A is caused by B based only on a correlation between these two variables (Bleske-Rechek et al., 2015; Rodriguez et al. 2016), failing to acknowledge the possibility of third variables that could be driving the relationship (Klaczynski et al., 1997; Shah et al., 2017).

Very few studies have investigated this bias in relation to graphical depictions of data. While Fansher et al. (2022c) reported that bar graphs and line graphs do not lead to increases in causality bias, Xiong and colleagues (2019) reported that scatterplots and bar graphs increased causal inferences compared to line graphs. These inconsistent findings could be partly due to methodological differences between studies; much research has shown that basic graphical elements and data features impact graph perception. In the current study, we specifically focused on scatterplots and investigated whether the presence of outliers in the data differentially impacted endorsement of correlational or causal inferences.

Outliers & Scatterplots

Visual outliers are features that deviate from a local distribution of features and their saliency is a function of their deviation from the remaining set (Rosenholtz, 1999; Hochstein et al., 2018). It is well established that the visual system can rapidly identify outliers (Treisman & Gelade, 1980; Wolfe, 1994). Furthermore, recent work has shown that outliers are represented with high resolution (Avcı & Boduroglu, 2020) and are obligatorily processed along with remaining items in a scene and cannot be totally disregarded (Cant & Xu, 2020). Despite these findings on processing of visual outliers, there is limited research translating and integrating these perceptual findings into the graph processing realm. Specifically, how presence of outliers influence scatterplot processing is still not fully understood.

It is known that in scatterplots, viewers can extract correlation magnitude with moderate accuracy, and these estimations are impacted by various visual features of the scatterplot. Early studies have typically asked participants to intuitively estimate the association strength between the variables depicted in scatterplots (e.g. Bobko & Karren, 1979; Cleveland, Diaconis & McGill, 1982; Cleveland & McGill, 1984; Meyer, Taieb & Flasher, 1997); a smaller subset of studies asked participants to discriminate between two simultaneously presented scatterplots whose certain properties were manipulated (e.g. Doherty, Anderson, Angott & Klopfer, 2007; Pollack, 1960; Rensink & Baldridge, 2010; Rensink, 2017). Regardless of the methodology used, studies

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

have typically shown that viewers' estimates are not veridical (e.g. Cleveland et al., 1982; Meyer et al., 1997), with estimates showing greater deviations from the actual for weaker than stronger correlations (e.g. Bobko & Karren, 1979; Cleveland et al., 1982; Correll & Heer, 2017; Meyer et al., 1997). A number of these earlier studies have also investigated how viewers are influenced by the presence of outliers in scatterplots while estimating associations between variables. These studies have shown that people typically disregard or minimize the impact of outlier(s) when making estimates (Bobko & Karren, 1979; Meyer & Shinar, 1992; Meyer et al., 1997; Rensink, 2010; Wainer & Thissen, 1979; for a more recent account see Ciccione et al., 2023).

Unlike these studies focusing on estimating correlation magnitude, a smaller set of studies have investigated how outliers impact extracted trends. For instance, Correll & Heer (2017) demonstrated that viewers excluded outlier clusters while estimating trend-lines. However, when the impact of single outliers on trend-line estimates were investigated, there was no evidence that they were excluded (Oral & Boduroglu, 2022). In addition to methodological details that vary across these studies, this seeming contradiction between the exclusion of outlier clusters and the inclusion of outlier points in trend estimates may be partly due to the perceptual difference between a cluster and a single outlier point. Specifically, the cluster may have been processed as a separate entity from the remaining items and could have been therefore excluded from trend estimates. While these studies speak to the processing of outliers in scatterplots, they do not address how outliers impact reasoning about data presented in scatterplots.

Present Study

In three experiments, we specifically investigated whether the presence of outliers in scatterplots influenced people's tendency to differentially endorse correlational and causal statements. Perception research suggests that people exclude distinct outliers as they summarize other remaining visual elements (Avcı & Boduroglu, 2021; Epstein et al., 2020). On the other hand, trend-line estimate studies suggest single outliers are not excluded from trend-line estimates. Even if outliers were initially excluded as trends are summarized, they could nevertheless impact subsequent visual reasoning processes. By visual reasoning processes we refer to processes related to the comprehension of the data. Specifically, whether people conclude there is a correlational or causal relationship between variables depicted. In all three experiments, we presented participants with a series of scatterplots; half of these scatterplots had an outlier that differed from the remaining data spatially on both axes (outlier+) but was trend-consistent. In Experiment 1 and 2, the outlier+ and no-outlier displays were matched on overall correlation magnitude. This meant that the major cluster in outlier+ displays (i.e. the outlier excluded) in fact depicted a weaker relationship than the matching no-outlier display. these experiments were identical except for the magnitude of the relationships depicted. Experiment 1 and 2 were identical in all regards except that in Experiment 1, scatterplots depicted moderate (.4-.7) and in Experiment 2, scatterplots depicted weaker relationships (.2-.4) (for example stimuli see Figure 1). In Experiment 3, we matched the magnitude of the relationship between the outlier-excluded cluster in outlier+ and the no-outlier scatterplots; outliers in the outlier+ scatterplots further strengthened an existing moderate-to-strong relationship (see Figure 2). In each experiment, after viewing each scatterplot, participants were presented with one of two statements, either a correlational statement (X and Y change together) or a causal statement (X changes Y) and were asked to rate their agreement on a 0-100 scale. The exact phrasing of these statements was determined after extensive piloting.

If participants were *not* excluding outliers while reasoning about the data presented in scatterplots, then there should be no difference in endorsement for correlational and causal statements in Experiment 1 and 2, because in these two experiments, the outlier+ and no-outlier displays were matched in magnitude. On the other hand, in Experiment 3, given outlier+ scatterplots depicted a stronger relationship than no-outlier scatterplots, and that outliers amplified an already moderate relationship, one would expect there to be greater support for correlational statements for outlier+ compared to no-outlier scatterplots.

If participants were excluding outliers during visual reasoning, then this could influence perceived correlational magnitudes; consequently, in the first two experiments for no-outlier scatterplots, correlational statements should be endorsed more strongly than for outlier+ scatterplots because the former would likely be perceived as depicting stronger relationships. We expected this to be particularly true for Experiment 1 where moderate relationships were depicted. In Experiment 2, since weak relationships were predicted, we thought there may be greater error in perceived correlational magnitudes (Rensink, 2017) and outlier presence may have less of an impact on endorsement of correlational statements. In Experiment 3, exclusion of outliers during the reasoning process should result in perception of similar magnitude relationships across the outlier+ and no-outlier displays; if this were the case, the endorsement for correlational statements should be similar.

While the literature does not allow us to make clear predictions regarding endorsement for causality statements, we expected people to make stronger endorsement for causality when the perceived relationship was stronger than weaker. Simply put, if it is thought that A and B are not related, then it would be unlikely that A and B would be thought of as causally related. This anticipation should result in stronger endorsement of causality in Experiment 1 (moderate relationships) than in Experiment 2 (weak relationships). With regards to the impact of outlier presence in scatterplots, whether they would impact perceived causality could vary as a function of whether they are excluded (or not). If outliers were excluded from the reasoning process, then causal statements could be endorsed more in the no-outlier than in the outlier+ condition in

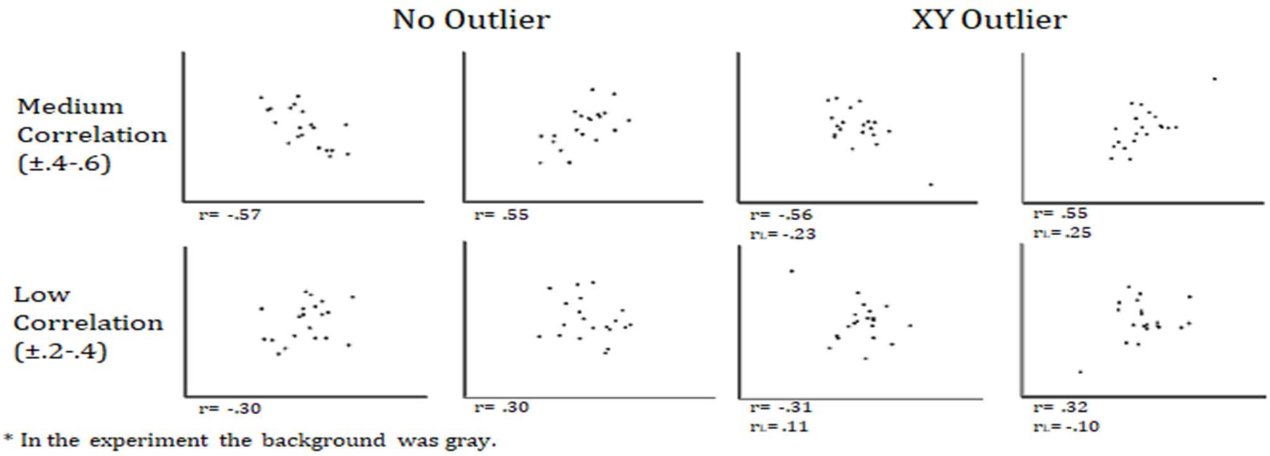

Figure 1: The top and bottom row presents example stimuli from Experiment 1 and 2, respectively. For the outlier+ graphs, the correlation magnitude with and without the outliers are both indicated.

Experiment 1; similar levels of causality endorsement would be expected across the two conditions in Experiment 3.

Given the similarity of the three experiments, the method and results of these three experiments are be presented together.

Method

Participants

In all three experiments, Bogazici University undergraduates, with normal or corrected-to-normal vision, participated in an online experiment in return for course credit. Based on a priori determined exclusion criteria, participants failing the attention checks, those who gave the same rating on more than 80 % of the trials, and whose ratings and study completion times exceeded that of the sample by 3 SD's were not included in subsequent analyses (26, 14 and 8 participants in Experiment 1-3, respectively). This left us with 106 participants (59 female, 44 male, 3 Other) in Experiment 1, 114 participants (64 female, 49 male, 1 Other) in Experiment 2 and 108 participants (60 female, 40 male, 8 Other) in Experiment 3. Target sample size was determined based on Xiong et al. (2019).

Scatterplot task

In all three experiments, participants completed the same scatterplot task. The task was programmed in PsyToolKit. We first describe the general structure of the task and then highlight experiment-specific variations.

The experiment started with a very brief familiarization phase where participants were shown an example scatterplot and introduced to graph elements (the x and y axis, the dots etc.). Participants were told that on each trial they would be presented with a scatterplot and a statement underneath. Upon inspecting the data presented in the scatterplot, they were to indicate how much they agreed with that particular statement using a 0-100 slider scale (anchors: I disagree vs. I agree; cursor left flushed). They were told to pay attention to the statements and that these statements would vary.

On each trial, participants were centrally presented with a scatterplot depicting 20 black points on a gray background; to minimize impact of prior beliefs, axis were not labelled. Underneath each scatterplot, there was one of two statements: "X changes Y" (causal statement) or "X and Y change together/X and Y jointly change" (correlational statement). For each experiment, 20 scatterplots were created, half depicting a positive and the other half depicting a negative relationship. Half of these were outlier+ scatterplots, yoked to no-outlier scatterplots. Each scatterplot was presented twice in random order, once coupled with the correlational and the other time coupled with the causal statement.

In outlier+ scatterplots, the spatial location of the outlier point was determined based on the trend of the data, ensuring that the outlier was trend-consistent. This meant that for positive trends the outlier was presented in either the lower left or the upper right corners of the remaining data cloud. The reverse was true for negative trends. The position of the outlier point was determined so that the corresponding value deviated at least by 2.5 SDs (max 3SDs) from the remaining data cloud on both the X and Y axis. In Experiment 1 & 2, the magnitude of the correlations for outlier+ and no-outlier scatterplots were matched. The only difference between these two experiments was that in Experiment 1 scatterplots depicted moderate relationships (average r = .65, with correlation magnitudes ranging between .42 - .69) and Experiment 2 depicted weak relationships (average r = .30, with correlation magnitudes ranging between .21 - .39). In Experiment 3, the global correlation magnitude in the no-outlier and the outlier+ was not matched. Instead, we created 19 point scatterplots and for the no-outlier condition, we added another dot around the centroid of the data cloud; for outlier+ scatterplots, we added an additional trend-consistent outlier point that further amplified the trend depicted by the 19 points in the data cloud. This resulted in no-outlier scatterplots that depicted moderate relationships (r =.5 - .7)

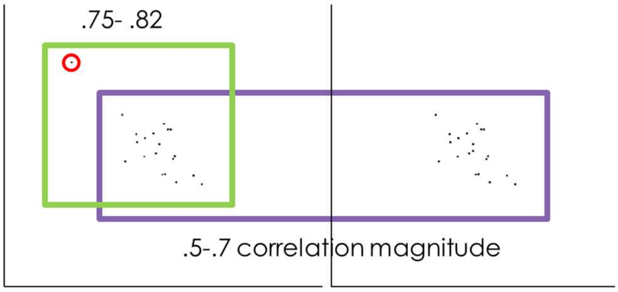

Figure 2: Example stimuli from Experiment 3. The left panel represent the trend-consistent outlier condition, and the right panel represents the no outlier condition. The outlier is circled for emphasis. In both cases (purple outlined), the main data cluster revealed a correlation of identical magnitude, but in the outlier condition, the outlier increased the strength of association.

and outlier+ scatterplots that depicted stronger relationships (r=.75 - .82), (see Figure 2).

Individual Differences Measures

Upon completion of the scatterplot task, in Experiment 1 and 2, participants also completed graph literacy and a scientific reasoning scales. To measure graph literacy, we used a 7-item subscale on scatterplot processing from the Visualization Literacy Assessment Test (Lee et al., 2017). To measure scientific reasoning, we used the 11 item Scientific Reasoning Scale (Drummond & Fischhoff, 2015). Performance on both these scales were highly and negatively skewed and because of restricted range problems we could not utilize these data as individual difference measures. Because of that, in Experiment 3, these scales were omitted.

Results

In all three experiments, for each participant we calculated the average endorsement for correlational and causal statements for outlier+ and no-outlier scatterplots. Table 1 presents the average endorsement for correlational and causal statements across the three experiments, and the comparison of the no-outlier and the outlier+ conditions. For each experiment, we separately carried out a 2 (Statement type: correlational, causal) X2 (Scatterplot type: no-outlier, outlier+) repeated measures ANOVA on the endorsement ratings.

In Experiment 1, the analyses revealed a main effect of scatterplot type (F(1,105)=24.53, p<.001,$\eta_p{}^2$ =.189) that was qualified by a scatterplot type X statement type interaction (F(1,105)=7.178, p<.008,$\eta_p{}^2$=.064). This pattern was due to participants giving higher ratings to no-outlier displays. More specifically, while participants gave much stronger endorsement to correlational statements in the no-outlier condition than in the outlier+ condition (d=.36), the same was not true for the causal sentences (d=.13). This suggested that participants may have been at least partially excluding the outlier during their visual reasoning process. Furthermore,

Table 1: Endorsement of correlational and causal statements

Statement	Expt.	No-Outlier M (SD)	Outlier+ M (SD)	Cohen's d
	Expt 1	53.6 (18.6)	46.6(20.0)	.36
Correlational	Expt 2	50.2 (21.1)	44.6 (21.6)	.26
	Expt 3	52.0 (20.9)	56.9 (20.3)	.24
	Expt 1	50.2 (19.4)	47.7(20.0)	.13
Causal	Expt 2	42.7 (22.0)	38.9 (22.1)	.17
	Expt 3	50.5 (21.9)	57.8 (21.9)	.33

the higher ratings for correlational statements may reflect a certain level of awareness that moderate correlations between variables do not necessarily imply a causal link; given that this particular sample had both high scientific reasoning and graph literacy this may not be surprising. Furthermore, there was a moderate relationship between endorsement for correlational statements and the correlation magnitude depicted in the scatterplots (r=.48, p=.034).

In Experiment 2, the same 2X2 ANOVA analyses revealed a main effect of statement type (F(1,113)=9.31, p<.002,$\eta_p{}^2$=.08) and a weaker yet significant main effect of scatterplot type (F(1,113)=23.84, p<.001,$\eta_p{}^2$=.02). The interaction did not reach significance, F<1. These results partially replicate Experiment 1 in that no-outlier scatterplots (M=46.5) received higher ratings than outlier+ (41.8) regardless of statement type. This suggests that participants may had been excluding, at least partially, the outlier as they reasoned about the data presented in the scatterplots, because as in Experiment 1, in Experiment 2, no-outlier displays depicted stronger relationships. Critically, this experiment also demonstrated that when weaker relationships were presented in scatterplots, endorsements of causality was lower (47.4 vs. 40.8), consistent with our expectations of a general awareness that unrelated (or weakly related) variables are less likely to be causally related.

In the third experiment, unlike in the first two, instead of equating the overall correlation magnitude, we had equated the correlation represented by the 19 dots in the outlier+ (i.e. excluding the outlier) and that represented in the no-outlier condition. In this case, the presence of a trend-consistent outlier amplified the local moderate relationship. If participants were excluding outliers regardless, then in this experiment, endorsement of statements should not have varied across conditions. Contrary to these expectations, in Experiment 3, outlier+ (57.4) scatterplots received higher endorsement than no-outlier (51.3) scatterplots, (F (1,107) =32.72, p<.001, $\eta_p{}^2$=.23). The statement type X scatterplot type interaction did not reach significance, (F (1,107) = 2.58, p=.11, $\eta_p{}^2$=.02). Nevertheless, to follow-up on our a priori expectations, our inspection of the endorsement of correlational and causal statements revealed an interesting pattern. There seemed to be a moderate impact of outlier presence on endorsement of causal statements (d=.33) and this effect was almost as strong as the effect of outliers on correlational statements (d=.36) in Experiment 1. We believe that these findings may suggest that when outliers further amplify an existing moderate trend, participants may be more likely to exhibit causality bias, misinterpreting correlational relationships as causal.

Discussion

These three experiments are among the first to investigate the impact of outliers on how people perceive both the correlational and causal relationships between variables depicted in scatterplots. Bringing together insights from visual perception, graph processing and reasoning, we believe our work offers an initial yet unique perspective as to when and why participants may exhibit a causality bias when viewing sets of correlational data. We specifically demonstrate that only when outliers enhance moderate relationships, participants are more likely to infer causality. However, in other conditions, participants refrained from making strong causal inferences. Thus, the impact of the outlier seems to be dependent on what is conveyed by the broader correlational data. Specifically, only when outliers increase the correlation magnitude depicted in the scatterplot, it leads to meaningful increases in the endorsement of correlational and causal relationships. Our findings go beyond recent studies that have investigated the relationship between graphical depictions of data and causality bias (e.g. Fansher et al., 2020c; Xiong et al., 2017) because those work do not directly address the issue of outliers. We argue that systematic investigations into the effect of outliers on reasoning is worthwhile because research has shown that people can focus on aberrations in data patterns especially when trying to challenge scientific truths (e.g. for the climate change context see Ranney & Clark, 2016). Furthermore, belief in pseudoscience is often associated with illusions of causality (Torres et al, 2020). If outliers in certain patterns enhance moderate relationships, this may lead people to misinterpret trends, and show unwarranted yet stronger belief in the data. If these data confirm with existing worldviews, this could further lead to confirmation and/or myside-type biases.

Due to restricted range problems we were not able to investigate the impact of graph literacy and scientific reasoning on scatterplot processing. Nevertheless, we must note that our particular sample of college students had high graph literacy and scientific reasoning scores. Taking this into account, the three experiments revealed that, individuals with high graph literacy and scientific reasoning, can accurately process data presented in scatterplots: For instance, in the first two experiments when presented relationships were weak or moderate, participants were not misguided by outlier values. They were able to disregard the impact of these outlier data points. People were also sensitive to the magnitude of the relationship depicted: endorsement of correlational statements increased as a function of actual correlation magnitude in the scatterplot. Comparison of endorsement given to both correlational and causal statements across Experiment 2 and 3 are consistent with this interpretation. Also, when scatterplots depicted weaker relationships, causal statements received lower endorsement (see Table 1).

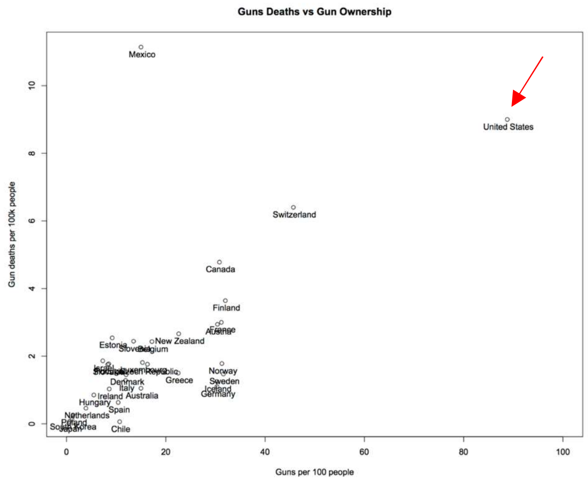

Figure 3: the relationship between gun ownership per 100 people and gun deaths per 100K people across countries.

It is possible that our choice to use no-context as we presented these scatterplots, may have influenced the results and possibly made it easier to ignore outliers in Experiment 1 and 2. If context were to be provided, participants' vulnerability to top-down factors could have influenced their endorsement for correlation and causal statements. Specifically, if an outlier point represented a familiar case, then its impact may be much harder to ignore, even if the data set represented a relatively weak relationship. We believe Figure 3 illustrates this possibility.

Figure 3 shows the relationship between gun ownership and gun deaths per 100K people across countries (figure and data from: http://mark.reid.name/blog/gun-deaths-vs-gun ownership.html). The red arrow was added to highlight the data point representing the United States. When the data from the US is excluded, the correlation magnitude of the remaining data is approximately around .40; if the US data is included, the correlation surpasses .60. As one inspects this data, once a salient and publicly discussed case is recognized, it might be much harder to ignore its impact. What remains unknown is whether in these cases people disproportionately exhibit causality bias. Given the importance of and challenges associated with everyday reasoning (Shah et al., 2019), future work needs to replicate and extend our current findings to investigate the impact of familiarity and plausibility of outliers on causality bias.

References

Avci, B., & Boduroglu, A. (2021). Contributions of ensemble perception to outlier representation precision. *Attention, Perception, & Psychophysics*, *83*, 1141-1151.

Beattie, V., & Jones, M. (2008). Corporate reporting using graphs: a review and synthesis. *Journal of Accounting Literature*, *27*, 71-110.

Bleske-Rechek, A., Morrison, K. M., & Heidtke, L. D. (2015). Causal inference from descriptions of experimental and non-experimental research: public understanding of correlation-versus causation. *The Journal of General Psychology,142*(1),48–70.

Bobko, P., & Karren, R. (1979). The perception of Pearson product moment correlations from bivariate scatterplots. *Personnel Psychology*, *32*(2), 313-325.

Cant, J. S., & Xu, Y. (2020). One bad apple spoils the whole bushel: The neural basis of outlier processing. *NeuroImage*, *211*, 116629.

Chatterjee, A., Nardi, C., Oberije, C., & Lambin, P. (2021). Knowledge graphs for covid-19: An exploratory review of the current landscape. *Journal of Personalized Medicine*, *11*(4), 300.

Ciccione, L., Dehaene, G., & Dehaene, S. (2023). Outlier detection and rejection in scatterplots: Do outliers influence intuitive statistical judgments? *Journal of Experimental Psychology: Human Perception and Performance*, 49(1), 129-144.

Cleveland, W. S., & Cleveland, W. S. (1985). *The elements of graphing data* (Vol. 2). Monterey, CA: Wadsworth Advanced Books and Software.

Cleveland, W. S., Diaconis, P., & McGill, R. (1982). Variables on scatterplots look more highly correlated when the scales are increased. *Science*, *216*(4550), 1138-1141.

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. Journal of the American statistical association, 79(387), 531-554.

Correll, M., & Heer, J. (2017, May). Regression by eye: Estimating trends in bivariate visualizations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 1387-1396).

Doherty, M. E., Anderson, R. B., Angott, A. M., & Klopfer, D. S. (2007). The perception of scatterplots. *Perception & Psychophysics*, *69*(7), 1261-1272.

Dragicevic, P., & Jansen, Y. (2017). Blinded with science or informed by charts? A replication study. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 781-790.

Drummond, C., & Fischhoff, B. (2017). Development and validation of the scientific reasoning scale. *Journal of Behavioral Decision Making*, *30*(1), 26-38.

Epstein, M. L., Quilty-Dunn, J., Mandelbaum, E., & Emmanouil, T. A. (2020). The outlier paradox: The role of iterative ensemble coding in discounting outliers. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(11), 1267.

Fagerlin, A., Zikmund-Fisher, B. J., & Ubel, P. A. (2011). Helping patients decide: ten steps to better risk communication. *Journal of the National Cancer Institute*, *103*(19), 1436-1443.

Fansher, M., Adkins, T. J., Lewis, R. L., Boduroglu, A., Lalwani, P., Quirk, M., ... & Jonides, J. (2022b). How well do ordinary Americans forecast the growth of COVID-19? *Memory & Cognition*, *50*(7), 1363-1380.

Fansher, M., Adkins, T. J., Lalwani, P., Boduroglu, A., Carlson, M., Quirk, M., ... & Jonides, J. (2022a). Icon arrays reduce concern over COVID-19 vaccine side effects: a randomized control study. *Cognitive Research: Principles and Implications*, *7*(1), 1-9.

Fansher, M., Adkins, T. J., & Shah, P. (2022c). Graphs do not lead people to infer causation from correlation. *Journal of Experimental Psychology: Applied*, *28*(2), 314.

Galesic, M., & Garcia-Retamero, R. (2011). Graph literacy: A cross-cultural comparison. *Medical Decision Making*, *31*(3), 444-457.

Klaczynski, P. A., Gordon, D. H., & Fauth, J. (1997). Goal-oriented critical reasoning and individual differences in critical reasoning biases. *Journal of Educational Psychology*, *89*(3), 470-485

Lee, S., Kim, S. H., & Kwon, B. C. (2017). Vlat: Development of a visualization literacy assessment test. *IEEE transactions on Visualization and Vomputer Graphics*, *23*(1), 551-560.

Lipkus, I. M., & Hollands, J. G. (1999). The visual communication of risk. *JNCI monographs*, *1999*(25), 149-163.

Merkl-Davies, D. M., & Brennan, N. M. (2007). Discretionary disclosure strategies in corporate narratives: incremental information or impression management?. *Journal of Accounting Literature*, *27*, 116-196.

Meyer, J., Taieb, M., & Flascher, I. (1997). Correlation estimates as perceptual judgments. *Journal of Experimental Psychology: Applied*, *3*(1), 3-14.

Oral & Boduroglu (2022). Outlier processing in scatterplots. *unpublished manuscript.*

Pollack, I. (1960). Identification of visual correlational scatterplots. *Journal of Experimental Psychology*, 59(6), 351-360.

Ranney, M. A., & Clark, D. (2016). Climate change conceptual change: Scientific information can transform attitudes. *Topics in Cognitive Ccience*, *8*(1), 49-75.

Rensink, R. A. (2017). The nature of correlation perception in scatterplots. *Psychonomic Bulletin & Review*, *24*, 776-797.

Rensink, R., & Baldridge, G. (2010). The visual perception of correlation in scatterplots. *Journal of Vision*, *10*(7), 975-975.

Rodriguez, F., Ng, A., & Shah, P. (2016). Do college students notice errors in evidence when critically evaluating research findings?. *Journal on Excellence in College Teaching*, *27*(3), 63-78.

Romano, A., Sotis, C., Dominioni, G., & Guidi, S. (2020). The scale of COVID-19 graphs affects understanding, attitudes, and policy preferences. *Health Economics*, *29*(11), 1482-1494.

Rosenholtz, R. (1999). A simple saliency model predicts a number of motion popout phenomena. *Vision Research*, *39*(19), 3157-3163.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational psychology review*, *14*, 47-69.

Shah, P., Michal, A., Ibrahim, A., Rhodes, R., & Rodriguez, F. (2017). What makes everyday scientific reasoning so challenging? In *Psychology of Learning and Motivation* (Vol. 66, pp. 251-299). Academic Press.

Tal, A., & Wansink, B. (2016). Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy. *Public Understanding of Science*, *25*(1), 117-125.

Tal, A. (2015). Looks like science, must be true: graphs and the halo of scientific truth. *Jury Expert*, *27*, 12.

Torres, M.N., Barberia, I & Rodriguez-Ferreiro, J. (202). Causal illusion as acognitive basis of pseudoscientific beliefs. *British Journal of Psychology*, 840-852.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97-136.

Tumen, C. & Boduroglu, A. (2022). Judgmental Time Series Forecasting: A systematic analysis of graph format and trend type. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44, No. 44).

Wainer, H., & Thissen, D. (1979). On the robustness of a class of naive estimators. Applied Psychological Measurement, 3(4), 543-551.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202-238.

Xiong, C., Shapiro, J., Hullman, J., & Franconeri, S. (2019). Illusion of causality in visualized data. *IEEE Transactions on Visualization and Computer Graphics*, *26*(1), 853-862.

Zacks & Franconeri (2021). Reading the pandemic. Editor's Blog. Science https://www.science.org/content/blog-post/read-pandemic-data

Acknowledgments

I would like to acknowledge contributions from Emre Oral and Burcu Avcı in early phases of this work.