How Prediction Visualization Improves Bar Charts Interpretation and Data Recall?

By

Minh Ngoc Le

* * * * * * * * *

Submitted in partial fulfillment
of the requirements for
Honors in the Department of Computer Science

UNION COLLEGE

June, 2020

Abstract

LE, MINH NGOC How Prediction Visualization Improves Bar Charts Interpretation and Data Recall?. Department of Computer Science, June, 2020.

ADVISOR: Kristina Striegnitz

In this paper, I presented an empirical study that analyzed the effectiveness of incorporating users' prior knowledge in visualization interaction. I introduced a design space for bar charts that facilitate two popular elicitation tasks, prediction and self-explanation, allowing users to actively engage and externalize their internal representations. The study was administered as a controlled, online experiment and was designed to evaluate how visually predicting and self-explaining the gap between one's expectation and the actual data can support chart comprehension and improve data recall. The results show that participants who were prompted to do the elicitation tasks outperformed the control group in recalling data, both in absolute and relative terms. However, these effects do not apply for misleading bar charts, and the size of learning gains varies across different spatial visualization ability.

Contents

List of Figures

List of Tables

1 INTRODUCTION

Graphs are a powerful tool to depict quantitative information and assist people in understanding data. The formal, theoretical frameworks for visualization such as graph grammars and transformation have been extensively developed since the late 1990s [8, 38, 39]. Recently, research in information visualization has shifted their focus towards the role of interactive visualization in human cognitive activities [25]. Although interactive tasks such as data retrieval, data filtering, and view manipulation are commonly used, most visualizations are still impersonal and do not incorporate people's internal representations. Internal representations are mental models of our knowledge, background, beliefs, and are critical in the learning process. Studies in cognitive science find that people develop internal representations as they view or create visualizations, and external depictions cannot entirely replace internal visualizations [25, 28, 39].

To target the lack of integrating internal representations into current visualizations, research has started to experiment and evaluate the use of elicitation tasks to connect people's internal knowledge and the external representations of data. Most existing works focus on more complex visualizations in facilitation for scientific investigation [11, 12, 35, 37]. However, given the prevalent use of visualization in education and popular media, more studies have recognized the importance of giving the general public access to such tools [9, 25] and attempted to design interaction interface to allow users to externalize their knowledge and expectation [5, 22]. New York Times' "You Draw It" series is an example of interactive visualization made available to the public. The tool prompts readers to draw lines reflecting their prediction before showing the correct data [21]. Also using line charts, Kim et al. (2018) take a step further and formally design a study to learn the effects of eliciting prior knowledge on cognition. Their results suggest that by asking viewers to predict and explain the data can significantly improve data interpretation and recall even when they have little expertise on the topics [22]. Given the great potential benefits of incorporating internal representations in visualization, existing research has not explored such effects in other types of graphs such as bar charts, pie charts, or maps. Another missing aspect from current research is misleading charts - in this age of data explosion, misleading graphics are prevalent in news and media to bias readers' view of complex and controversial issues. Whether introducing intervention can have a positive impact on misleading charts is not yet studied.

I expand on prior work with two main contributions. First, I focus my study on bar charts, evaluating how effectively elicitation tasks can improve viewers' graph understanding and cognition. Second, I test whether such effects persist in misleading charts. The results show that externalizing viewers' prior knowledge leads to better data recall and comprehension, especially for people who are asked to predict the data beforehand. These effects, however, are insignificant in misleading bar charts.

2 RELATED WORKS

2.1 Internal Representations

Studies have looked into graph comprehension, learning how people interpret charts and why it is easier for people to perceive and understand data in pictorial and graphic form [31, 23]. Simply present graphs and pictures may not lead to a better understanding of data compared to the original formats, one needs to also consider factors such as cognitive load, mental effort, and internal representations [17]. Using a top-down perspective, research finds that the interplay between internal and external representations, in particular, can greatly enhance the cognitive outcome of visualization [25, 28, 39, 22]. Broadly, internal representations can be defined as the mental images and symbolic descriptions of what people perceive from the world with the visual sensation that is stored in the brain [32]. Specifically for InfoVis, Liu and Stasko (2010) define internal representation as a functional, structural, behavioral analog representation to external interactive visualizations.

The strong focus on developing external visualization reflects the desire to replace the need for internal visualization. However, studies have suggested that external visualizations have to be coupled with internal visualization, and it is important to foster their connection for better learning outcomes [11, 25, 37]. For example, engaging in mental animation before or instead of external animated visualization helps viewers understand the content better [13]. Externalizing one's internal representation or self-constructing visualization is also found to be effective in enhancing people's understanding of data [7, 14].

2.2 Elicitation Techniques

The formal, theoretical frameworks for studies in visualization including graph grammars and transformation have been extensively developed since the late 1990s [8, 38, 39]. Initially, these methods mostly focus on the external appearance of data such as the accuracy of displaying quantitative information [3], graph layout, and navigation [15]. As visualization evolved, user interactions are found to be as important to visualization experience as the presentation of data [41, 24]. Interaction techniques often use WIMP (windows, icons, mouse, pointer) actions to perform selection, exploration, filtering, or more advanced tasks such as navigation and visual comparison [40, 36, 10]. However, interactive visualization still falls short in connecting people's internal, mental faculties with external representations of data. Cognitive and psychology studies have suggested that people always have an internal representation of the dataset while viewing the graphics. Thus, external representations should involve internal mental models to produce more effective visualizations [7, 25, 37].

With the popularization of visualization, especially in education and popular media, studies in InfoVis have started to expand their scope and develop applications that serve broader communities. They also shift focus from analytic to communicative visuals [16, 19, 33]. To incorporating internal representations in information visualization, research has developed and assessed the effectiveness of several elicitation techniques such as prediction and self-explanation. Cox (1999) compares the cognitive difference between presented visualization and self-constructed visualization. He found that by actively constructing the external representation, people excel in problem-solving. This is because self-construction facilitates better idea organization, mental animation, information re-ordering, and self-explanation effects. Prediction is a type of self-constructed visualization and is shown in later literature to effectively assist analytical reasoning [22, 34]. Viewing the gap between one's prediction and the actual data also implicitly generate self-learning as people try to repair their knowledge accordingly [22, 27].

Self-explaining data and graphics is also an effective way to engage people's mental models in learning activities. Self-explanation prompts people to reflect on both their prior knowledge and the presented data, thus, leads to better knowledge integration and understanding [4]. Comparing text and diagrams, Ainsworth et al. (2003) find that diagrams promote significantly larger self-explanation effects and more effective learning compared to text.

Specifically, in terms of charts, Kim et al. (2017) carry out a controlled experiment for line charts. They find that participants who are involved in prediction and self-explanation tasks outperform in data recall and comprehension. The combination of both prediction and self-explanation is shown to be the most effective. They also confirm previous findings that information visualization is more effective in presenting data than text. The effect is also larger for people who have low familiarity with the data. This is because (1) it is easier for them to repair their mental models given the limited knowledge they have, and (2) the curiosity effect that increases the desire to seek the knowledge to fill the gap [22, 26].

2.3 Misleading Data Visualization

Data visualization is a powerful communication tool that can deliver messages in persuasive and comprehensible ways, however, it can also be easily misused or misinterpreted. In the field of journalism and media, some intentionally use graphs to shape convincing narratives and bias readers' opinions on important matters. This problem has long been discussed in literature and research. Visualization deceptions can occur at two levels: the chart level where people read and interpret the charts incorrectly, and the message level where people get the wrong message from the graph. At the message level, there are two types of deceptions: message exaggeration/understatement and message reversal [30]. How may these deceptions emerge? The classic ”How to Lie with Statistics” [18] or the more recent ”How to Lie with Charts” [20] show various ways statistical and numerical communication may lead to misinformation. One can choose to obfuscate data and obscure reality with extreme simplification or biased data selection. Alternatively, performing axis manipulation such as y-axis truncation, broken axes, or logarithmic scale often result in message exaggeration or understatement [18, 20, 2, 29, 6].

Some initial attempts have been made on how to help people overcome the deceptive nature of misleading charts. One way is to pair deceptive graphs with correct, explanatory text. This method, however, only works with certain demographics and overall results still suggest that deceptive practices can still alter people's ability to read and understand the data [29]. Given the effectiveness of elicitation tasks in enhancing visual cognition, existing literature lacks investigation into the use of such methods for misleading visualizations. Therefore, in my thesis, I also explore the potential use of elicitation tasks for misleading visualizations.

3 METHODOLOGY

3.1 Elicitation Techniques

Based on previous literature, I incorporated two elicitation techniques to externalize internal representations in information visualization. The specification is as follows:

1. **Prediction**: prompt users to predict the data before they view the correct ones. Studies have shown that actively constructing external visualizations and predicting data lead to viewers' deeper understanding of the dataset. [7, 22, 27, 34].

2. **Self-explanation**: prompt users to reflect and write down the reasons why there are differences between their expectation/prediction and the data. Self-explanation facilitates the process of identifying the discrepancy between one's prior knowledge and actual data which can lead to better cognitive outcomes [4, 22].

Given the elicitation mechanism, I devised four different conditions for the experiment. All the data were presented using visualizations in term of bar charts:

- **Baseline**: users only view and examine the actual data, no intervention is introduced.

- **Predict-Only**: users are presented with a bar chart missing one column and are asked to predict the missing data (by dragging up/down the bar to adjust the height). After the prediction, the users will be able to view and compare their predictions side-by-side to the correct data.

- **Explain-Only**: users are immediately shown the actual data. After observing the correct charts, users are asked to provide some explanations about the gap between their expectations and the actual data by typing into a text box.

- **Predict-Explain**: include both elicitation tasks. After making predictions on the missing columns, they will observe the correct data side-by-side with their predictions and input some explanations into a text box.

Due to sample size restriction, I only tested one elicitation condition for misleading charts - I chose a combination of prediction and self-explanation which is found to be the most effective for line charts [22]. This results in 6 different conditions: 4 for normal charts (Baseline-Normal, Predict-Only-Normal, Explain-Only-Normal, Predict-Explain-Normal) and 2 for misleading charts (Baseline-Misleading, Predict-Explain-Misleading).

3.2 Misleading Charts

I employed one of the most popular ways to create deceptive effects on bar graphs: truncating the y-axis. Instead of the standard baseline of 0 and a maximum of 100, the misleading bar graphs in the study have the y-axis started at a different level above 0 and ended below 100. Axis truncation can exaggerate small differences between the bars, leading to message exaggeration/understatement deception effects.

Figure 1 shows two of the charts used for the experiment, comparing the difference between normal and misleading charts.

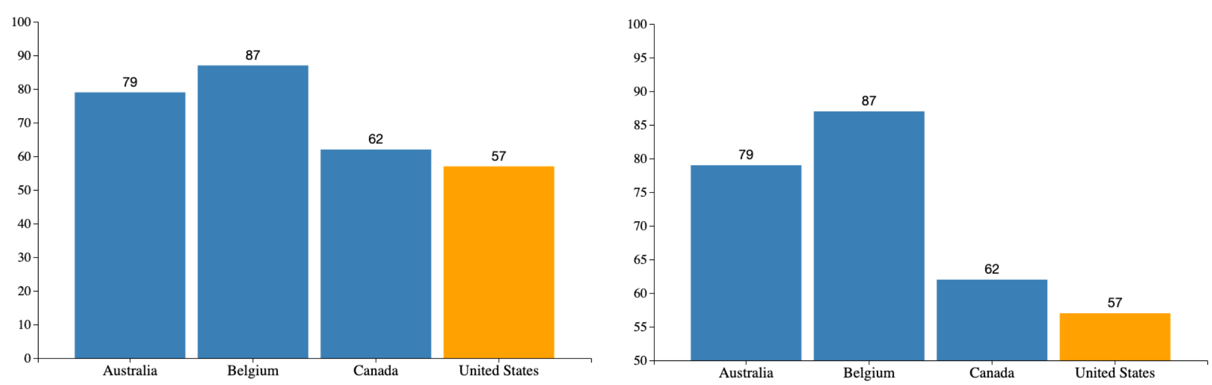

Figure 1: Comparison between a normal chart (left) and a misleading chart (right)

3.3 Hypotheses

- **H1**: I expect that participants in predict conditions (Predict-Only, Predict-Explain) will be able to recall data more accurately than those in the baseline condition. Literature suggests that prediction can implicitly help people extract their internal representations and improve data cognition. Here, prediction can be especially helpful because people can visually and directly observe the difference between their expectations and the correct data without much effort [7, 4, 22, 34].

- **H2**: I expect participants in Explain-Only condition to recall data better than those in the baseline condition. Self-explanation can increase graph comprehension and learning outcome [4, 1, 22].

- **H3**: I expect the use of elicitation tasks will improve viewers' performance for misleading charts.

3.4 Choice of Datasets

I choose three different topics to create bar graphs for the experiments:

- Voting turnout in the most recent election for developed countries (Australia, Belgium, Canada, United States)

- United Stated federal spending by major categories in 2015

- United States mortality rate by causes in 2015

Post experiment survey on perceived familiarity reveals that more than two-thirds of the participants report they are neither clearly familiar nor unfamiliar with all the three topics. Therefore, results interpretation can factor out the effects of familiarity without subjecting to significant biases.

All conditions used identical knowledge contents, but misleading charts presented the data using different y-axis. The charts were in the same format, consisting of numbers in percentage and four categorical dimensions, resulting in four data points in each graph. This format is common for bar graphs in science and media. To make it more challenging for people to capture the value and trend, the bars are placed in alphabetical order, i.e., they are do not follow ascending or descending order.

3.5 Procedure

Figure 2 illustrates the design of the experiment. I employed the between-group design: participants are randomly placed into one of the six conditions. After reviewing general information about the study and confirming their consent to participate, participants completed a series of tasks depending on the condition

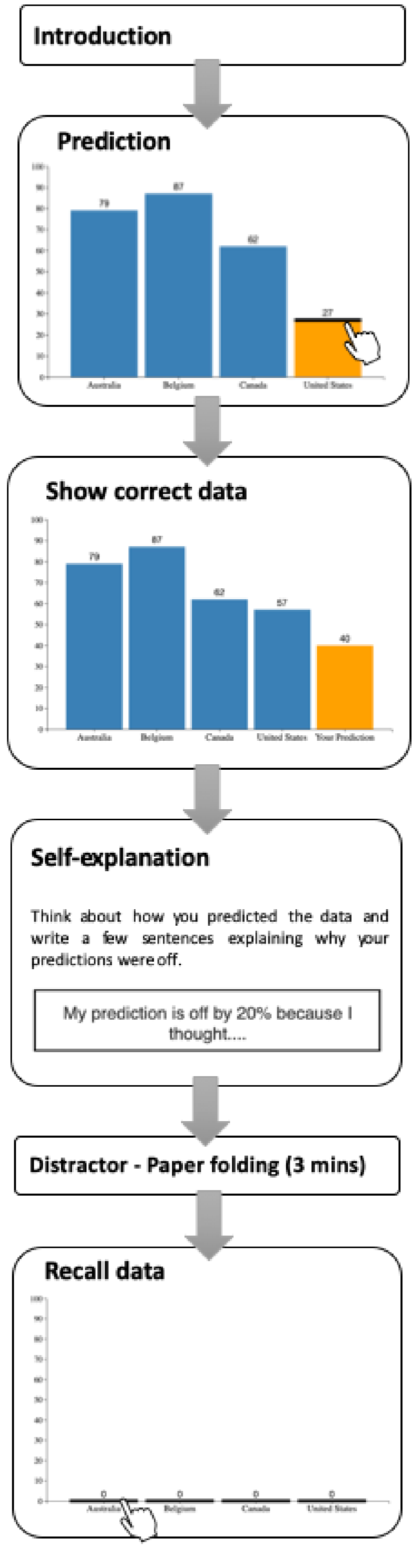

Figure 2: The procedure of the online experiment: (1) introduction, (2) make prediction by dragging the bar to adjust the height, (3) view the correct charts side-by-side with previous predictions, (4) write down reflections on the data, (5) paper folding task, and (6) recall data.

they were in. For prediction conditions (Predict-Only-Normal, Predict-Explain-Normal, Predict-Explain-Misleading), each graph has exactly one missing column, the participants were prompted to predict a value for this column by adjusting the height of the bar. After drawing their predictions for all three graphs, they viewed the correct bars and their expectation side-by-side. Their predictions were in a different color to distinguish from actual data. In other conditions where there is no prediction task, the participants were shown the correct graphs immediately after the introduction page.

On the same page where participants examining the data, for self-explanation conditions (Explain-Only-Normal, Predict-Explain-Normal, Predict-Explain-Misleading), participants were asked to write a few sentences in a text box placed below the graphs to help they understand the graphs better. Two sample questions were provided as guidelines, asking how much their expectations were off from the actual data and why they may have been wrong.

Next, all participants were presented with 10 paper folding tasks that they needed to complete within three minutes. This serves both as a distractor task and as a way to collect information on spatial capability. Previous studies have shown that there is a positive correlation between people's spatial abilities and how effective they can use internal representations and recall visualizations [14, 12, 14, 22].

Finally, participants were asked to recall all the data points for the three graphs they previously saw. Here, both normal and misleading conditions were shown standard graphs with the y-scale starts at 0. This poses a particular challenge for those in misleading conditions since they were exposed to graphs with various scales earlier in the study.

Before submitting the experiment, the participants completed a basic demographic questionnaire, asking about their age, gender, ethnicity, region, and educational attainment. The survey also asked people to indicate their familiarity with graphs and the three topics used to compute the graphs in this experiment.

4 RESULT

4.1 Data Preliminaries

I recruited 120 workers on Amazon Mechanical Turk (AMT), each condition had on average 20 responses. The average time to complete the assignment was 20 minutes with no major variation between conditions. The workers must reside in the United States and be above the age of 18. Each participant was compensated with $3 upon completion. There is no significant difference in demographic distribution (Appendix). I excluded 4 observations that are outliers from the sample.

4.2 Analysis Approach

4.2.1 Dependent Variables

I analyzed how accurately participants recalled the data by using two main error indicators: absolute and relative error. Absolute error reflects the ability to recall the numerical values and equals to the absolute sum of all the differences between one's recalled and correct data points. The relative error shows how well participants can recall the height, or trend, of the bars relative to one another. The maximum relative error value is 18 (3 graphs x 6 pairs/graph).

Figure 3 shows an example of a recalled graph. Here, the absolute error is $\mid 50 - 79 \mid + \mid 87 - 80 \mid + \mid 62 - 63 \mid + \mid 40 - 57 \mid = 54$. The relative height of Australia and Canada is wrong, so the relative error is 1 in this case.

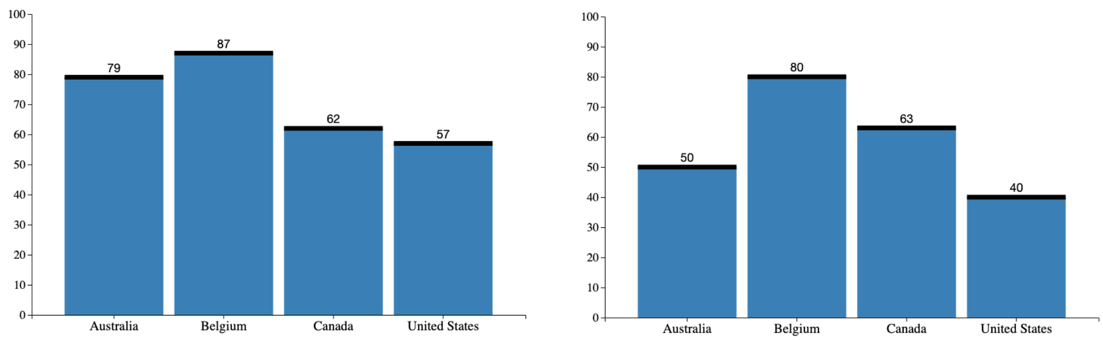

Figure 3: Comparison between a correct chart (left) and a recalled chart (right)

4.2.2 Regression Model

I followed Kim et al. (2017) and used a mixed effects model to evaluate the results. R's lme4 package produces the estimated coefficients and p-value for all regression models [22]. The model specification is as follows:

$$Error_i = \beta_0 + \beta_1 \times ExplainOnly_i + \beta_2 \times PredictOnly_i + \beta_3 \times PredictExplain_i + \beta_4 \times Spatial_i + \alpha_i + \gamma_i + \epsilon_i$$

where $Error_i$ is either absolute or relative error specified above; $ExplainOnly_i$, $PredictOnly_i$, $PredictExplain_i$ are binary variables that equals 1 if participant i is placed in that condition; $Spatial_i$ is the total correct answers from the paper folding task, centered by its mean to obtain the average spatial ability of the participant; α_i and γ_i are random effects on the participant's id and region, respectively. Since the misleading scenario only has two conditions, the model above will only include $PredictExplain_i$ variable. The omitted variable in both normal and misleading scenarios is the normal, no intervention condition - this is used as the reference point for the model. Therefore, the estimated coefficients on $ExplainOnly_i$, $PredictOnly_i$, $PredictExplain_i$ record how well the elicitation conditions perform compare to the baseline condition.

All regression results are reported using the 95% confidence intervals. The estimated coefficients are in terms of percentage - the same unit of the datasets used to compute the graphs in the experiment.

4.3 Normal Charts

4.3.1 Absolute Error

The intervention of elicitation tasks improves data recall in all conditions compared to the baseline (Figure 4). More specifically, participants in Predict-only see the largest decrease in absolute error, by -72.67 percent[1] (t = -7.54, p $<$ 0.001), followed by Explain-Only (-45.46, t = -4.77, p $<$ 0.001) and Predict-Explain (-40.91, t = -4.30, p $<$ 0.001). There is no significant difference between the three treatment conditions. These results show that interaction with internal representations, achieved through prediction and self-explanation, can lead to better absolute data recall - a proxy for visual cognition and memorability.

Spatial ability also plays a role in participants' ability to recall. Every additional correct answer in the paper folding test is associated with -3.6 percent decrease in absolute errors (t = -2.27, p $<$ .05). This further confirms previous findings that there is a positive correlation between spatial ability and effective visual cognition [11, 14, 22].

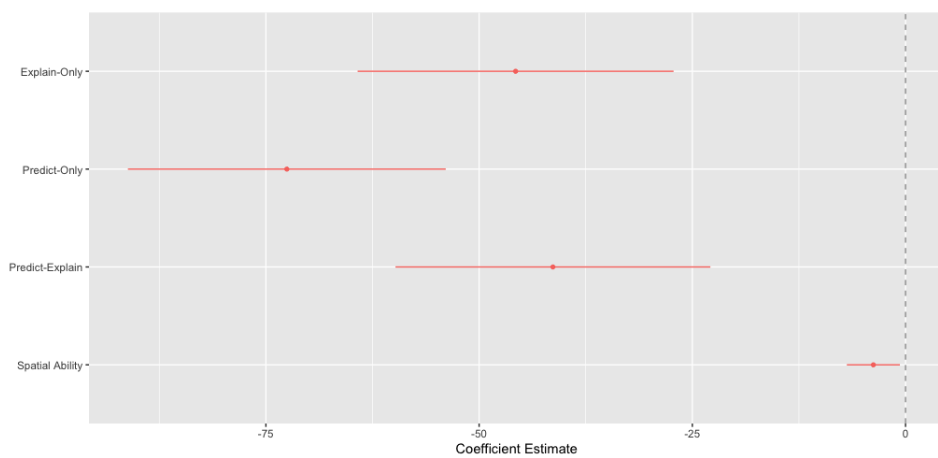

Figure 4: Estimated coefficients for absolute errors in normal charts

[1] Unit of measurement used in the three datasets, not a percentage change

4.3.2 Relative Error

In terms of relative error, only Predict-Explain condition has a significant decrease compared to the baseline condition (Figure 5). Being in Predict-Explain lower the error by -1.78 on average relative to the baseline ($t = -2.04$, $p < 0.01$). Spatial visualization ability continues to show an impact on recall error. Additional correct answer in the paper folding task reduce the relative error by roughly -0.46 ($t = -3.73$, $p < 0.001$).

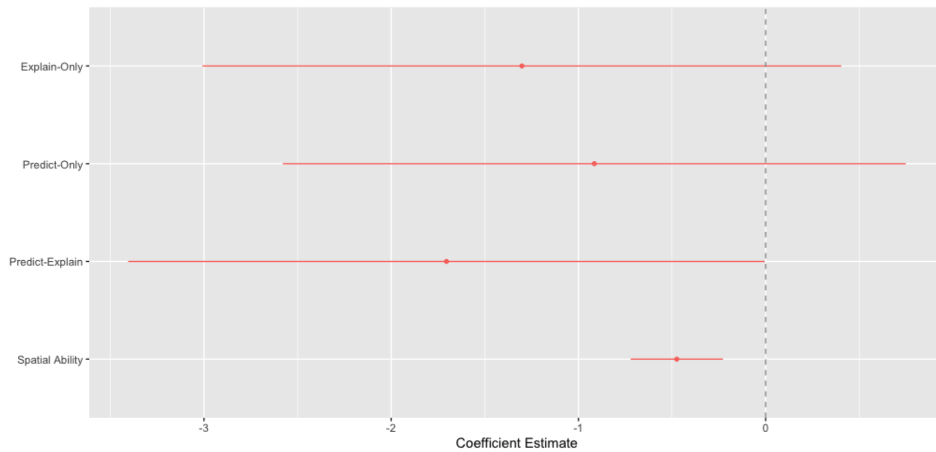

Figure 5: Estimated coefficients for relative errors in normal charts

4.4 Misleading Charts

First, it would be helpful to compare the baseline condition of normal and misleading charts. As expected, participants in the misleading condition reported 28.7 percent higher in absolute error on average ($t = 3.87$, $p < 0.01$). Misleading charts use various tactics such as improper scaling or cherry-picking data to distort the perception of data. In this experiment, I use y-axis truncation - starting the bar at a baseline other than zero - to exaggerate small differences between categories. Truncated graphs are shown to have a persistent subjective impact across visualization designs and increase estimation error, especially in terms of individual values. This is because the visual impression of the size of the differences between bars has a strong influence on the perceived values, thus, truncated graphs can create a deceptive effect that hinders readers from interpreting the data properly and leads to incorrect visual memories [2, 6].

Unlike in normal charts, I observe no significant difference in both absolute and relative errors between the Baseline-Misleading and the Predict-Explain-Misleading condition (Figure 6-7). As mentioned above, misleading charts can distort data so much that readers are prone to misinterpret the data and underlying trends. Thus, even with the intervention of prediction and self-explanation tasks that are supposed to help participants actively engage with the data, participants still could not effectively overcome the biases created by the deceptive visualizations. This is not an unusual result since there have been multiple unsuccessful attempts to improve people's perception of misleading graphs. For example, people are susceptible to distortion techniques even when deceptive graphs are accompanied with accurate information in the form of explanatory text, and as a result, are misinformed about the presented data [29].

Spatial visualization ability continues to show positive effects on data recall. Each additional correct answer in the paper folding task lowers absolute error by approximately -4.03 percent (t = -2.05, $p < 0.01$) and relative error by -0.61 (t = -2.91, $p < 0.01$). This result once again further confirms the idea that spatial ability can lead to better visual perception and recall, even in the context of misleading charts.

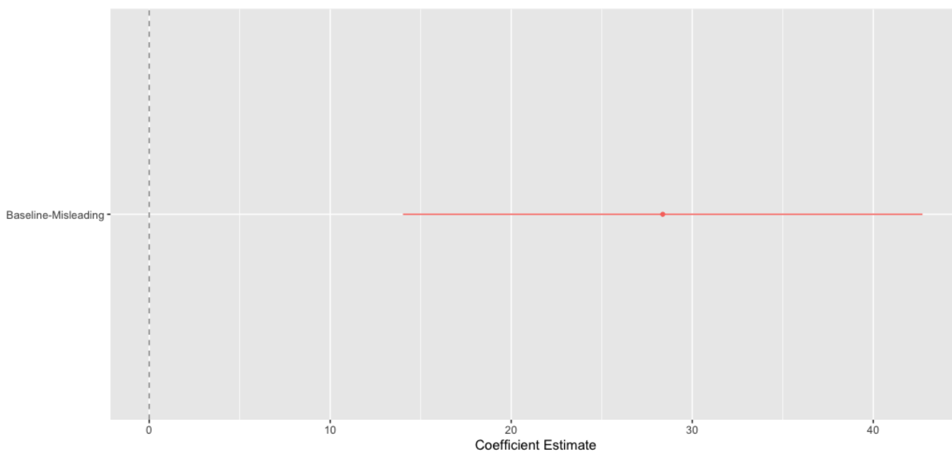

Figure 6: Comparison between Normal and Misleading baseline conditions

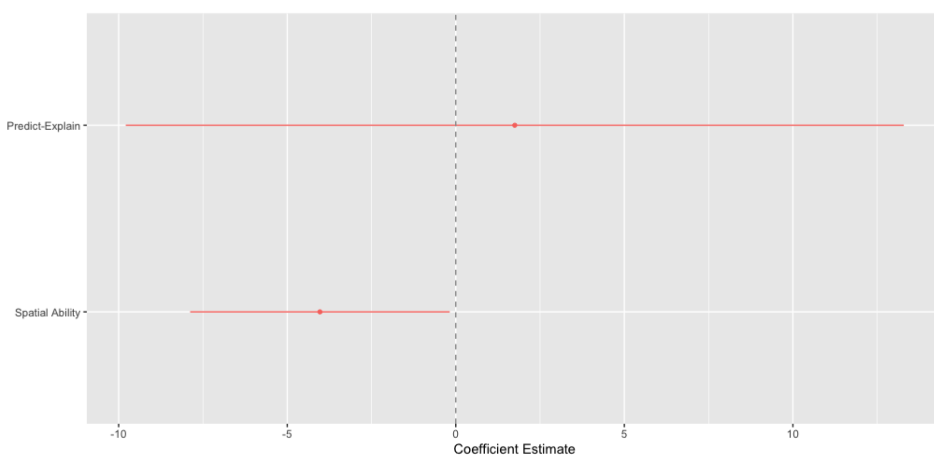

Figure 7: Estimated coefficients for absolute errors in misleading charts

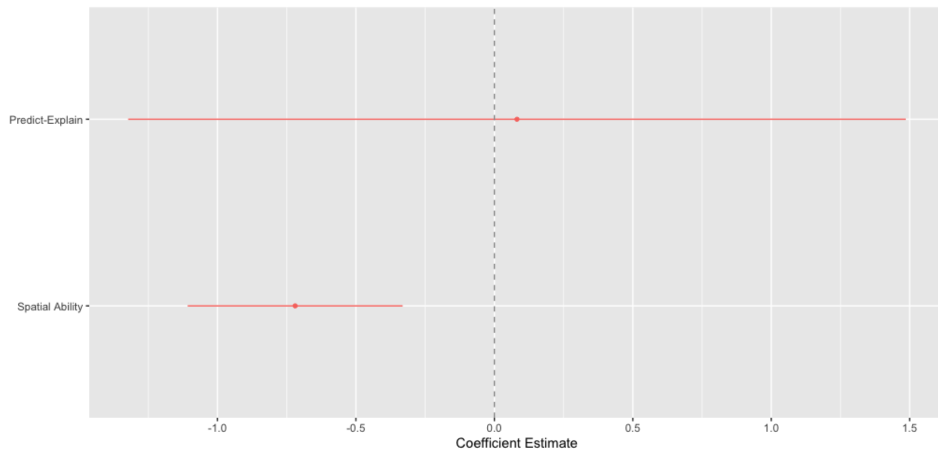

Figure 8: Estimated coefficients for relative errors in misleading charts

4.5 Learning Gain

Given the significant impact of spatial visualization ability on data recall, I took a step further and investigate whether spatial ability matters to how much one can potentially gain doing the elicitation tasks. I compared the absolute errors obtained during the prediction step when participants were asked to draw the last column of each graph and those obtained during the recall phase. Three conditions were included: Predict-Only-Normal, Predict-Explain-Normal, and Predict-Explain-Misleading. As a proxy for learning gain, I computed the percentage change in absolute errors before and after the participants observed the actual data:

$$LearningGain_i = \frac{PredictionErrors_i - RecallErrors_i}{PredictionErrors_i} \times 100$$

I divided the sample into 5 different groups by participants' scores in the paper folding tasks which varied from 0-10. Using this approach, I assumed that participants only differed in spatial ability and were roughly the same in other aspects, even though they were exposed to different elicitation conditions. A simple linear regression model is used to estimate how each spatial group performed, with the lowest spatial group (score 0-1) as the omitted reference. I also included educational level as an explanatory variable since people with higher education may have gained more skills to effectively read and learn from graphs. The model specification is as follows:

$$LearningGain_i = \beta_0 + \beta_1 \times SpatialGroup2_i + \beta_2 \times SpatialGroup3_i + \beta_3 \times SpatialGroup4_i + \beta_4 \times SpatialGroup5_i + \beta_5 \times Education_i + \epsilon_i$$

The regression results confirm the role of spatial ability in learning visualization. Only two groups with the highest paper folding scores observe a significant change in absolute errors. Specifically, participants who scored 8-10 in the paper folding task recalled data 37.14 percent better than those who scored 1 or below (t = 2.13, $p < 0.05$). Those who got 6-7 answers right also saw an equivalent change, recalling data 35.41 percent more accurately (t = 2.01, $p < 0.05$). Participants who only got half or less of the paper folding questions did not perform better than the reference group. Education level also has a positive correlation on the magnitude of learning gain, with higher education level leads to roughly 20 percent gain in recall accuracy.

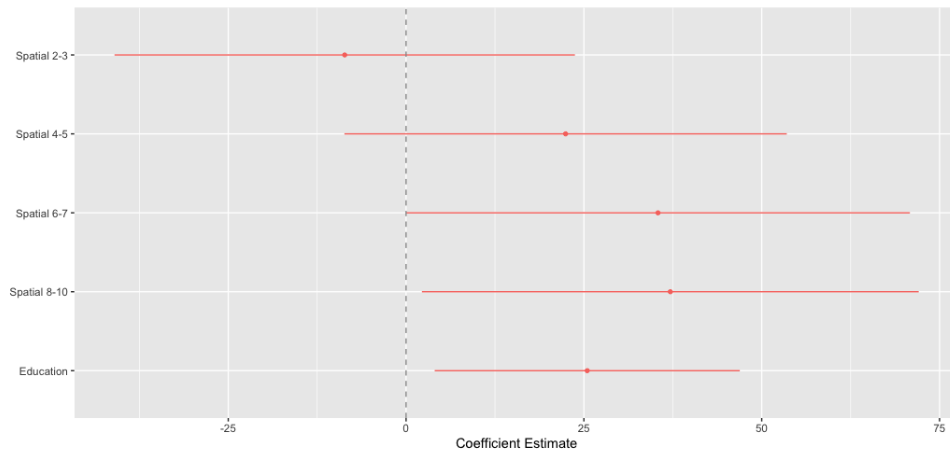

Figure 9: Estimated coefficients for learning gains

5 DISCUSSION

5.1 Contribution

This paper extends prior research on the use of elicitation methods in enhancing visual learning outcomes and the effect of misleading visualization on data perception. I introduce a design space to externalize people's internal representation such as their prior knowledge and expectations in bar graphs. This is achieved by asking the users to drag up and down a bar to set the height indicating their value predictions.

The results presented in this paper further supports the findings of Kim et al. (2017) that prompting users to either predict or self-explain graphs can significantly improve data recall. Similar to the previous experiment, though participants only predicted one data point for each graph, both Predict-Only and Predict-Explain conditions observed a decrease in absolute errors. In fact, prediction tasks are superior to explanation and the combination of prediction and explanation when it comes to recalling individual data points. This may suggest that participants likely based their estimations using the existing information on the graph, and thus, were engaged with the whole dataset [22]. In terms of recalling relative bar heights, Predict-Explain is the only condition that reports a significant change in the error. This may imply that the relative trend of the bars is harder to capture and thus, requires the user to engage more deeply with graphs. Therefore, by completing both prediction and explanation tasks, there is a higher chance that the participants recognized the correct ranking of the bars and recalled them correctly later on.

To further confirm the positive effects of elicitation tasks, prediction in particular, and learn about the role of spatial ability in data cognition, I also carried out a study on learning gains. I compared the absolute errors at the time of the prediction task and data recall, the results show that those with higher spatial visualization ability benefit more from doing the elicitation tasks. This may be crucial for future works since the designers of visualization would ideally want to achieve an equal learning effect across all groups.

My work is the first to incorporate elicitation tasks in misleading visualizations through the use of truncated bar graphs. Although I expected that these tasks would improve users' data recall, I did not observe any significant effect on either absolute or relative errors. This may be because deceptive visualizations can greatly distort users' understanding and perception of data. In the experiment, as participants were first exposed to a different scaling of the y-axis, they registered the data associated with such scale, leading to exaggeration or underestimation in their memories. The wrong perceived messages of the visualizations make it harder for them to recall the data points on a normal scaling chart later on.

5.2 Limitations and Future Work

My study has a number of limitations. First, the sample size is relatively small and may not be a reliable representation of the broader population. Since I did not carry out a preliminary survey to choose the datasets, I may not have picked the best ones for this experiment. I also deliberately picked datasets with only categorical values, thus, future work should evaluate the same techniques with other datasets and chart formats such as time series or stacked bar charts.

For misleading charts, I only ran the experiment on two conditions: the baseline and predict-explain. Future studies should incorporate all available elicitation techniques to fully assess their effects. Since misleading charts influence people's perception of data much differently from normal charts, there is also a need for new elicitation methods and interfaces to be developed.

Besides data recall, future works should consider studying other potential use of elicitation tasks such as addressing people's existing biases and improving their ability to detect and interpret the underlying messages.

6 CONCLUSION

Based on previous findings in visualization and cognitive science, I designed a study to evaluate the effectiveness of incorporating internal representations on users' visual cognition in bar charts. The results show that eliciting users' prior knowledge through prediction and self-explanation tasks is an effective way to improve chart comprehension and data recall. These effects, however, are not significant for misleading visualizations and vary greatly across different spatial ability groups. It is important that we continue to research on the practical use of interactive visualization to connect people with their mental models and improve their ability to overcome deception and misinformation.

References

[1] Shaaron Ainsworth and Andrea Th Loizou. “The effects of self-explaining when learning with text or diagrams”. In: *Cognitive science* 27.4 (2003), pp. 669–681.

[2] Alberto Cairo. “Graphics lies, misleading visuals”. In: *New Challenges for Data Design*. Springer, 2015, pp. 103–116.

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

[4] Michelene TH Chi et al. “Eliciting self-explanations improves understanding”. In: *Cognitive science* 18.3 (1994), pp. 439–477.

[5] In Kwon Choi et al. “Visual (dis) Confirmation: Validating Models and Hypotheses with Visualizations”. In: *2019 23rd International Conference in Information Visualization–Part II*. IEEE. 2019, pp. 116–121.

[6] Michael Correll, Enrico Bertini, and Steven Franconeri. “Truncating the Y-Axis: Threat or Menace?” In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, pp. 1–12.

[7] Richard Cox. “Representation construction, externalised cognition and individual differences”. In: *Learning and instruction* 9.4 (1999), pp. 343–363.

[8] Rozenberg Grzegorz. *Handbook Of Graph Grammars And Computing By Graph Transformations, Vol 2: Applications, Languages And Tools*. world Scientific, 1999.

[9] Zhan Guo. “Mind the map! The impact of transit maps on path choice in public transit”. In: *Transportation Research Part A: Policy and Practice* 45.7 (2011), pp. 625–639.

[10] Christopher G Healey and Brent M Dennis. “Interest driven navigation in visualization”. In: *IEEE transactions on visualization and computer graphics* 18.10 (2012), pp. 1744–1756.

[11] Mary Hegarty. “Diagrams in the mind and in the world: Relations between internal and external visualizations”. In: *International Conference on Theory and Application of Diagrams*. Springer. 2004, pp. 1–13.

[12] Mary Hegarty. “Dynamic visualizations and learning: Getting to the difficult questions”. In: *Learning and Instruction* 14.3 (2004), pp. 343–351.

[13] Mary Hegarty, Sarah Kriz, and Christina Cate. “The roles of mental animations and external animations in understanding mechanical systems”. In: *Cognition and instruction* 21.4 (2003), pp. 209–249.

[14] Mary Hegarty and Kathryn Steinhoff. “Individual differences in use of diagrams as external memory in mechanical reasoning”. In: *Learning and Individual differences* 9.1 (1997), pp. 19–42.

[15] Ivan Herman, Guy Melançon, and M Scott Marshall. “Graph visualization and navigation in information visualization: A survey”. In: *IEEE Transactions on visualization and computer graphics* 6.1 (2000), pp. 24–43.

[16] Jeremy Heyer, Nirmal Kumar Raveendranath, and Khairi Reda. “Pushing the (Visual) Narrative: the Effects of Prior Knowledge Elicitation in Provocative Topics”. In: (2020).

[17] Weidong Huang, Peter Eades, and Seok-Hee Hong. “Measuring effectiveness of graph visualizations: A cognitive load perspective”. In: *Information Visualization* 8.3 (2009), pp. 139–152.

[18] Darrell Huff. *How to lie with statistics*. WW Norton & Company, 1993.

[19] Jessica Hullman and Nick Diakopoulos. “Visualization rhetoric: Framing effects in narrative visualization”. In: *IEEE transactions on visualization and computer graphics* 17.12 (2011), pp. 2231–2240.

[20] Gerald E Jones. *How to lie with charts*. Sybex San Francisco, 1995.

[21] JOSH KATZ. “You Draw It: Just How Bad Is the Drug Overdose Epidemic?” In: *New York Times* (2017). URL: https://www.nytimes.com/interactive/2017/04/14/upshot/drug-overdose-epidemic-you-draw-it.html?mtrref=www.google.com&assetType=PAYWALL.

[22] Yea-Seul Kim, Katharina Reinecke, and Jessica Hullman. “Explaining the gap: Visualizing one’s predictions improves recall and comprehension of data”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2017, pp. 1375–1386.

[23] Stephen M Kosslyn. “Understanding charts and graphs”. In: *Applied cognitive psychology* 3.3 (1989), pp. 185–225.

[24] Shixia Liu et al. “A survey on information visualization: recent advances and challenges”. In: *The Visual Computer* 30.12 (2014), pp. 1373–1393.

[25] Zhicheng Liu and John Stasko. “Mental models, visual reasoning and interaction in information visualization: A top-down perspective”. In: *IEEE transactions on visualization and computer graphics* 16.6 (2010), pp. 999–1008.

[26] George Loewenstein. “The psychology of curiosity: A review and reinterpretation.” In: *Psychological bulletin* 116.1 (1994), p. 75.

[27] Danielle S McNamara. “iSTART: A Web-Based Tutor That Teaches Self-Explanation and Metacognitive Reading Strategies: Danielle S. McNamara, Tenaha O’Reilly, and Michael Rowe”. In: *Reading Comprehension Strategies*. Psychology Press, 2007, pp. 400–423.

[28] Donald A Norman. “Some observations on mental models”. In: *Mental models*. Psychology Press, 2014, pp. 15–22.

[29] Shaun O'Brien and Claire Lauer. "Testing the susceptibility of users to deceptive data visualizations when paired with explanatory text". In: *Proceedings of the 36th ACM International Conference on the Design of Communication*. 2018, pp. 1–8.

[30] Anshul Vikram Pandey et al. "How deceptive are deceptive visualizations? An empirical analysis of common distortion techniques". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2015, pp. 1469–1478.

[31] Steven Pinker. "A theory of graph comprehension". In: *Artificial intelligence and the future of testing* (1990), pp. 73–126.

[32] Zenon W Pylyshyn. "What the mind's eye tells the mind's brain: A critique of mental imagery." In: *Psychological bulletin* 80.1 (1973), p. 1.

[33] Edward Segel and Jeffrey Heer. "Narrative visualization: Telling stories with data". In: *IEEE transactions on visualization and computer graphics* 16.6 (2010), pp. 1139–1148.

[34] Elsbeth Stern, Carmela Aprea, and Hermann G Ebner. "Improving cross-content transfer in text processing by means of active graphical representation". In: *Learning and Instruction* 13.2 (2003), pp. 191–203.

[35] Mike Stieff. "Mental rotation and diagrammatic reasoning in science". In: *Learning and instruction* 17.2 (2007), pp. 219–234.

[36] Christian Tominski, Camilla Forsell, and Jimmy Johansson. "Interaction support for visual comparison inspired by natural behavior". In: *IEEE Transactions on Visualization and Computer Graphics* 18.12 (2012), pp. 2719–2728.

[37] J Gregory Trafton, Susan B Trickett, and Farilee E Mintz. "Connecting internal and external representations: Spatial transformations of scientific visualizations". In: *Foundations of Science* 10.1 (2005), pp. 89–106.

[38] Edward R Tufte. *The visual display of quantitative information*. Vol. 2. 2001.

[39] Edward R Tufte et al. *Visual explanations: images and quantities, evidence and narrative*. 1998.

[40] Yingcai Wu et al. "OpinionSeer: interactive visualization of hotel customer feedback". In: *IEEE transactions on visualization and computer graphics* 16.6 (2010), pp. 1109–1118.

[41] Xiaoru Yuan et al. "Intelligent graph layout using many users' input". In: *IEEE transactions on visualization and computer graphics* 18.12 (2012), pp. 2699–2708.

7 APPENDIX

Table 1: Perceived Familiarity of the Datasets

	Voting Turnout			National Spending			Mortality Rate		
Condition	Not	Somewhat	Very	Not	Somewhat	Very	Not	Somewhat	Very
Baseline-Normal	4	13	2	2	17	0	3	13	3
Predict-Only-Normal	7	15	1	5	18	0	9	12	2
Explain-Only-Normal	5	11	4	2	14	4	5	12	3
Predict-Explain-Normal	5	13	2	4	12	4	7	12	1
Baseline-Misleading	4	11	2	3	12	2	5	10	2
Predict-Explain-Misleading	2	13	2	0	13	4	2	13	2

Table 2: Gender Distribution - No difference across conditions ($p = 0.473$)

Condition	Female	Male
Baseline-Normal	12	7
Predict-Only-Normal	19	4
Explain-Only-Normal	12	8
Predict-Explain-Normal	12	8
Baseline-Misleading	13	4
Predict-Explain-Misleading	11	6

Table 3: Age Distribution - No difference across conditions ($p = 0.839$)

Condition	18-24	25-34	35-44	44-54	55+
Baseline-Normal	2	10	4	2	1
Predict-Only-Normal	2	10	5	3	3
Explain-Only-Normal	1	7	7	3	2
Predict-Explain-Normal	2	7	7	2	2
Baseline-Misleading	0	11	2	3	1
Predict-Explain-Misleading	2	7	6	1	1

Table 4: Ethnicity Distribution - No difference across conditions ($p = 0.527$)

Condition	White	Black	Other
Baseline-Normal	16	2	1
Predict-Only-Normal	16	2	5
Explain-Only-Normal	14	2	4
Predict-Explain-Normal	18	2	0
Baseline-Misleading	12	1	4
Predict-Explain-Misleading	8	1	8

Table 5: Education Distribution - No difference across conditions (p = 0.192)

Condition	High School	Bachelor	Graduate	Other
Baseline-Normal	3	8	7	1
Predict-Only-Normal	6	14	2	1
Explain-Only-Normal	2	11	6	1
Predict-Explain-Normal	12	5	3	0
Baseline-Misleading	2	14	1	0
Predict-Explain-Misleading	2	12	3	0

Table 6: Region Distribution - No difference across conditions (p = 0.192)

Condition	Midwest	Northeast	South	West
Baseline-Normal	8	3	7	1
Predict-Only-Normal	5	6	6	6
Explain-Only-Normal	4	9	4	3
Predict-Explain-Normal	3	6	6	5
Baseline-Misleading	4	4	4	5
Predict-Explain-Misleading	3	9	3	2

Table 7: Frequency of Using Graphs - No difference across conditions (p = 0.750)

Condition	Never	Rarely	Sometimes	Often	Very Often
Baseline-Normal	0	3	10	5	1
Predict-Only-Normal	1	7	5	7	3
Explain-Only-Normal	1	4	9	4	2
Predict-Explain-Normal	0	9	7	3	1
Baseline-Misleading	2	3	4	7	1
Predict-Explain-Misleading	0	4	7	3	3

Table 8: Descriptive Statistics

Variable	Obs	Mean	Std. Dev	Min	Max
absoluteError	116	141.49	93.02	7	386
relativeError	116	5.94	3.60	0	14
spatialAbility	116	4.38	2.72	0	10