

# TrueVisLies Study - Documentation

Graziano Blasilli & Marco Angelini

## Contents

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>Introduction</b>   | <b>1</b> |
| <b>2</b> | <b>Misleading Errors</b>                                      | <b>1</b> |
| 2.1      | Visualization design violation . . . . .                      | 1        |
| 2.2      | Reasoning Errors . . . . .                                    | 2        |
| <b>3</b> | <b>Visualization rhetoric</b>                                 | <b>3</b> |
| <b>4</b> | <b>Author Intent and Factors Behind the Misleading Effect</b> | <b>3</b> |

## 1 Introduction

This document provides the background knowledge needed to complete the user study. It summarizes the main categories of **misleading errors** that can occur in charts (Section 2), as well as the **rhetorical techniques** that visualization creators may use to create and present the visualizations, (Section 3), and the **author's intents** and factors behind the misleading effect of a visualization (Section 4), that consciously or not—to influence the reader interpretation.

## 2 Misleading Errors

This section describes common visualization mistakes and reasoning flaws that can cause a chart to convey an incorrect or biased message. These definitions are adapted from prior research [1] and include both visual design violations (such as truncated axes or inappropriate encodings) and reasoning errors (such as cherry-picking or misleading causal claims).

### 2.1 Visualization design violation

- **Truncated axis** A truncated axis does not start at zero. This can occur in both horizontal and vertical axes.
- **Dual axis** A dual axis is when there is more than 1 (vertical, typically) axis on the same chart (does not include 2 charts side-by-side).
- **Value as area/volume** Value as area/volume is when a numeric value is encoded as area (e.g., circles) or volume (e.g., cubes).

- **Inverted axis** Inverted axis is when the axis directions are opposite to what one would expect (e.g., cases increasing = down). This can occur in both horizontal and vertical axes.
- **Uneven binning** Uneven binning is when a chart's bins are arbitrarily defined or have uneven bin sizes.
- **Unclear encoding** Unclear encoding is when it is not possible to completely understand what is shown in a given chart due to issues such as poor image quality, overcomplicated representation, or simply not enough information to understand the chart, because of issues such as the legend being cut off.
- **Inappropriate encoding** Inappropriate encoding is when the chart's type (line chart, pie chart, etc.) is clearly not appropriate given the underlying data. For instance, showing a value change year-to-year with a stacked bar chart, as opposed to a simple or clustered one. Or a change in a value is plotted in a case where you would expect the absolute value instead.

## 2.2 Reasoning Errors

- **Cherry-picking** Cherry-picking is when the main conclusion of a chart and its accompanying text (title, annotation, caption) is consistent with the incomplete evidence presented but likely would not be generalizable with more representative evidence.
- **Setting an arbitrary threshold** Setting an arbitrary threshold is when a chart's accompanying text (title, annotation, caption) uses an arbitrary (i.e., not present in visualized data, not known, not justified) threshold to make judgments, inferences, and derive insights from the visualization. The threshold can be stated explicitly as a number or added as an annotation in a chart.
- **Causal inference** Causal inference is when a chart's accompanying text (title, annotation, caption) suggests a cause-and-effect relationship in an attempt to explain certain salient features of the chart and evaluate them. Causal relationships are typically evaluated either by themselves against an arbitrary (e.g., message author-defined) satisfactory threshold or against another inferred causal relationship. Causality inferred from a visualization can be especially misleading in cases when the data are cherry-picked.
- **Issues with data validity** A message has issues with data validity when the message's chart doesn't include explicit information about caveats related to data accuracy and data interpretation, or the message's text doesn't reflect existing caveats in their interpretation of the visualization.
- **Failure to account for statistical nuance** Failure to account for statistical nuance occurs when a message's chart or text makes a conclusion using inferences drawn without considering statistical nuances, such as the base rate fallacy.
- **Misrepresentation of scientific studies** Misrepresentation of scientific studies occurs when a chart's accompanying text (title, annotation, caption) relies upon pseudoscience to come to a conclusion.
- **Incorrect reading of chart** Incorrect reading of a chart is when a message's author cites a chart as evidence to a conclusion, but the author fails to correctly read the data presented by said chart.

### 3 Visualization rhetoric

This section introduces five types of rhetorical strategies identified in visualization research [2]. These strategies illustrate how choices in data selection, visual encoding, language, and interaction can influence viewers' interpretations.

- **Information Access Rhetoric** Refers to choices about what data to include, omit, filter, or aggregate. These choices shape what the viewer sees and which patterns or narratives are emphasized.
- **Provenance Rhetoric** Relates to how a visualization communicates trustworthiness, such as by citing sources, explaining methods, noting uncertainty, or including methodology or assumptions.
- **Mapping Rhetoric** Involves how data values are translated into visual features such as position, size, scale, metaphor, or color. This can create emphasis or shape perception beyond what the raw data conveys.
- **Linguistic-Based Rhetoric** Uses language elements like titles, labels, annotations, or captions to guide interpretation. Often includes metaphor, irony, rhetorical questions, or other narrative framing.
- **Procedural Rhetoric** Relates to interaction and navigation design, such as default views, filters, or animations. These features guide how users explore the data and can influence which conclusions they reach.

### 4 Author Intent and Factors Behind the Misleading Effect

- **Aesthetic-Driven Misrepresentation** The author prioritizes the visual appeal of the visualizations, giving them precedence over accuracy. These stylistic preferences, such as decorative choices, lead to misinterpretation, despite the author's intention not to mislead.
- **Bias Exploitation** The visualization is crafted to take advantage of known human visual perceptual or cognitive biases in order to push the reader to a particular interpretation.
- **Claim-Supporting Manipulation** The author shapes the visualization and its caption to highlight a desired pattern or hide counter-evidence. The goal is to persuade the reader to adopt a particular perspective, for example, by spreading disinformation about COVID-19.
- **Context Distortion** The author removes (or hides/omits) relevant context that is essential for accurate interpretation of the visualization (missing labels, explanations, legends). The goal is to persuade the reader towards a particular perspective.
- **Deliberate Reader Confusion** The author introduces unnecessary complexity, such as ambiguous encodings, to create confusion and make the reader more receptive to the author's narrative.
- **Lack of Visualization Literacy** The author is unfamiliar with best practices in visualization and inadvertently introduces misleading elements. For example, the author may choose inappropriate chart types or a design that inadvertently leads to misleading conclusions.

- **Selective Reporting** The author reports only partial information, such as, reporting partial data, focusing/highlighting on a data subset. The goal is to provide a skewed representation of the information, aiming at persuading the reader towards a particular perspective.
- **Space and Format Constraints** Constraints (such as, space constraints, image size/format) limit the design or presentation of the visualization. The author, to cope with these constraints, may incur information loss, which can lead to misrepresentation.
- **Unintentional Context Omission** Accidental omissions of relevant context that are essential for accurate interpretation of the visualization (missing labels, explanations, legends). The misleading purpose is not deliberate.

## References

- [1] M. Lisnic, C. Polychronis, A. Lex, and M. Kogan, “Misleading beyond visual tricks: How people actually lie with charts,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI ’23, (New York, NY, USA), Association for Computing Machinery, 2023.
- [2] J. Hullman and N. Diakopoulos, “Visualization rhetoric: Framing effects in narrative visualization,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2231–2240, 2011.