

## How Deceptive are Deceptive Visualizations?: An Empirical Analysis of Common Distortion Techniques

Pandey, A. V., Rall, K., Satterthwaite, M. L., Nov, O., & Bertini, E., (2015), *New York University Public Law and Legal Theory Working Papers*. 504.

### Introduction

- 近年、専門家でなくとも視覚化を作成できる簡単なソフトウェアのおかげで、視覚化の使用は確実に増加した
  - 視覚化によって、メッセージがより分かりやすく、説得力のあるものにできる場合、視覚化の表現は、作成者に簡単に誤用され、読み手に誤解される可能性がある
    - ◇ この問題は視覚的な表現に限定されず、数字や統計を用いたコミュニケーションの一般的な問題
  - データの視覚化が一般的に広まったために、誤用や誤解が人々に与える影響も増加したかもしれない
    - ◇ 例：メディアやTV、インターネット (Figure 1)



Figure 1. Some of the real-world data visualization examples which might lead to misinterpretation of message, hence to deception.

- 本研究では、読み手の騙される程度と、歪められた視覚化の影響と個人差の間に関連があるかどうかを検討する
  - 先行研究では、歪んだ視覚化の危険性は警告するものの、その影響力を評価することを目的とした実証研究は見られなかった
- 本研究の目的
  - ✓ メッセージの反転とメッセージの誇張/過小評価に焦点を当て、読み手を騙すような視覚化の影響を分類する
  - ✓ 実験的な環境で歪みの影響を再現するための例を提示する
    - ◇ よく見られる歪められた視覚化の表現手法をいくつか選択し、歪められた視覚化と、歪められていない視覚化を作成

- ◇ 個人差が騙されやすさに関連するかを調べるために、学歴、グラフの親密度 (familiarity)、画像処理能力 (visual ability) を測定

### Related Works

- 視覚化の研究者は、画像のエンコードが知覚レベルで情報をどのように歪ませるかを検討した
  - 特に、大きさや色、角度などが異なる表現をされたときに、データがどのように認識され、比較されるかを検討した研究
    - ◇ グラフの知覚に関する有名な実験 (Cleveland & McGill, 1984) では、いくつかのグラフ表現手法が、他より正確な量的情報の比較を導くことを示した
    - ◇ 色の選択によっては、多くの歪みが生じる (Brewer, 1997; Levkowitz & Herman, 1992)
    - ◇ 散布図と平行座標 (parallel coordinates) によって、相関の認識が異なる (Li, Martens, & Van Wijk, 2010; Rensink & Baldridge, 2010)
      - 相関を推定するように求められた場合、平行座標プロットにおいて正の相関を過小評価し、負の相関を過大評価する
    - ◇ 面積でなく、円の半径に数量を配置することは、グラフの歪みに繋がる (Kong, Heer, & Agrawala, 2010)
- これらの研究は主に視覚化効果に焦点を当て、知覚課題に基づいた
  - 本研究では、より日常的な場面を想定し、視覚化によるメッセージレベルの影響を検討する
  - よって、知覚課題ではなく、言語を用いて、オブジェクト間の数量を比較するような課題を用いる
    - ◇ グラフ理解のエンコードと解釈の段階で、歪みの影響が生じる可能性があるため、データの視覚化のメッセージレベルに焦点を当てた検討は重要

### Deception and Deceptive Visualizations

- 本研究における、歪められた視覚化の定義：
  - ✓ 騙す意図があってもなくても、実際のメッセージと異なるメッセージおよび構成要素についての信念を生成する可能性のあるグラフ表現
- 視覚化によって騙される可能性は、2つのレベルであり得る
  - グラフレベル: 読み手がグラフを誤って読み取る and/or 提示されたデータの見積もりを誤って処理する
    - ◇ 視覚化のエンコードのレベルで発生し、グラフを正しく読み取れるか否かで変化する
  - メッセージレベル: 読み手がメッセージを誤って解釈する
    - ◇ メッセージの解釈のレベルで発生し、メッセージおよび構成要素について誤

った信念を生み出す可能性がある

- 本研究では、メッセージレベルの影響にのみ焦点を当てる
  - ✓ **Message Exaggeration/Understatement** (メッセージの誇張/過小評価)
    - 提示された事実は誇張されるものの、歪められていない場合に生じる
      - ◇  $A > B$  というグラフにおいて A と B の数量を比較する場合、A は B よりも大きいという事実は提示されるものの、その範囲は誇張されている
    - この歪みは、「A は B よりどれくらい大きいと思いますか (How much)」といった質問に影響する
  - ✓ **Message Reversal** (メッセージの反転)
    - 読み手が事実を誤って解釈するよう、視覚化によって促される場合に生じる
      - ◇ A と B の数量を比較する場合、 $A > B$  というグラフにも関わらず、読み手は「A は B よりも小さい」というメッセージを知覚する
        - 実際のデータが示されていても、読み手は視覚化が歪んでいるために、誤ったメッセージを知覚する
    - この歪みは、「グラフは何を示していますか (What)」といった質問に影響する

## Study Rationale and Methods

### Selecting Distortion and Visualization Techniques

- ほとんどの視覚化は、歪んでいたり嘘をついたりする傾向がある (Jones, 2011)
  - しかし、歪み手法は固有のものであり、一つの手法が全ての視覚化に影響するわけではない
- 本研究では、広く使用される 4 つの歪み手法に焦点を当てる
  - 広く使用されている簡単な図表に焦点を当てるため、ヒートマップやネットワークマップのような複雑な図表は除外する

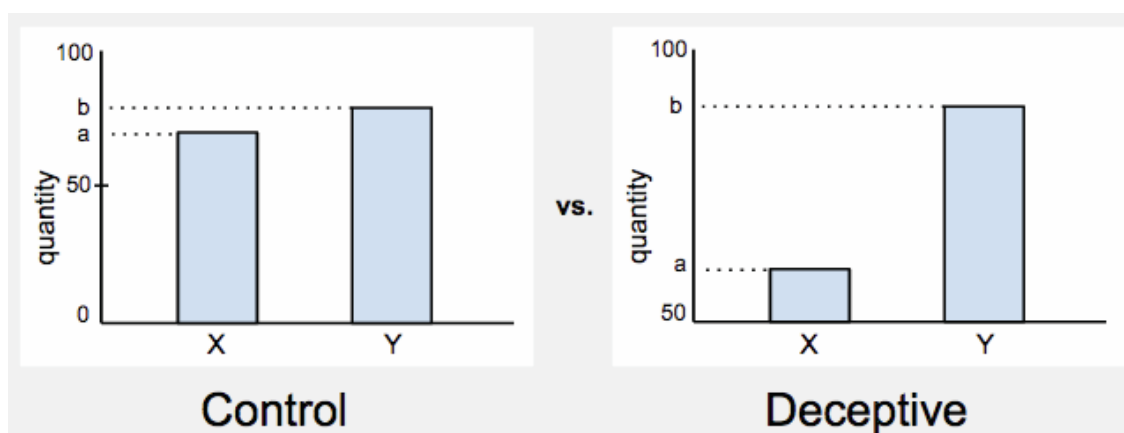


Figure 2. Illustration showing Truncated Axis distortion, which leads to message exaggeration/understatement type of deception.

### Creating Treatments

- ✓ **Truncated Axis** (軸の切り取り)
  - スケール上に表示される最小値と最大値を操作する(Figure 2)
    - ◇ このような軸の範囲の変更は、提示された数量の誇張もしくは過小評価に繋がり、「How much」形式の質問に対する読み手の反応に影響する
  - 軸ベースのグラフは全て、この種類の歪みの影響を受けやすい
  
- ✓ **Area as Quantity** (量としての面積)
  - 視覚化において量的データをサイズで表現するには慎重な設定が必要
    - ◇ データを面積で表すガイドラインはないが、データと面積が 1 対 1 になる設定において、最も歪みが少ないと考えられる (Tufte, 1983)
  - しかし、グラフ表現に影響する変数にデータが設定されることも、しばしばある
    - ◇ たとえば、データが円の半径に設定されている場合、Y は X より「かなり大きい」と判断されかねない (Figure 3)
  - 領域ベースのグラフは、この種類の歪みが発生しやすい
    - ◇ 統制条件のように、量的データが面積に設定されている場合、視覚化の影響は受けにくい

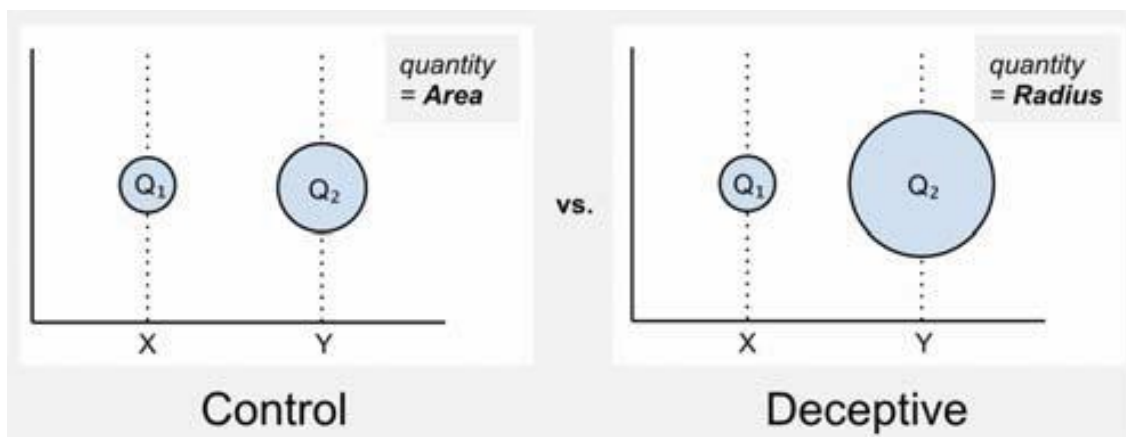


Figure 3. Illustration showing Area as Quantity distortion, which leads to message exaggeration/understatement type of deception.

- ✓ **Aspect Ratio** (アスペクト比)
  - アスペクト比により傾斜角度が変化し、量の増減の割合に影響する
    - ◇ 軸の目盛を広げると角度が歪み、統制条件よりも増加が緩やかに見える (Figure 4)
  - この種の歪みは、主に折れ線グラフで見られ、メッセージの誇張/過小評価を招く

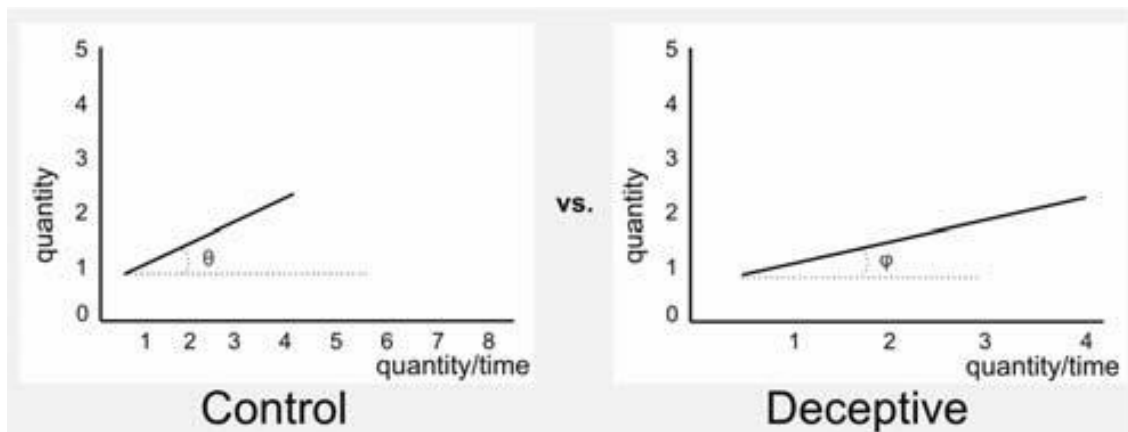


Figure 4. Illustration showing altered angle of the trend responsible for "rate of increase/decrease" due to Aspect Ratio distortion, which leads to message exaggeration/understatement type of deception.

✓ **Inverted Axis (軸の反転)**

- 人は方向と傾向を関連づける (Jones, 2011)
  - ◇ 上向きを増加、下向きを減少、右方向を進行、左方向を後退など
- これにより、軸の反転はメッセージの反転を生じさせる
- この種の歪みは、軸に依存するほとんど全ての視覚化テクニックに影響する
  - ◇ Figure 5 は、x 軸変数による y 軸変数の増加を示すが、右側の図は y 軸が反転しているために、減少している印象を与える

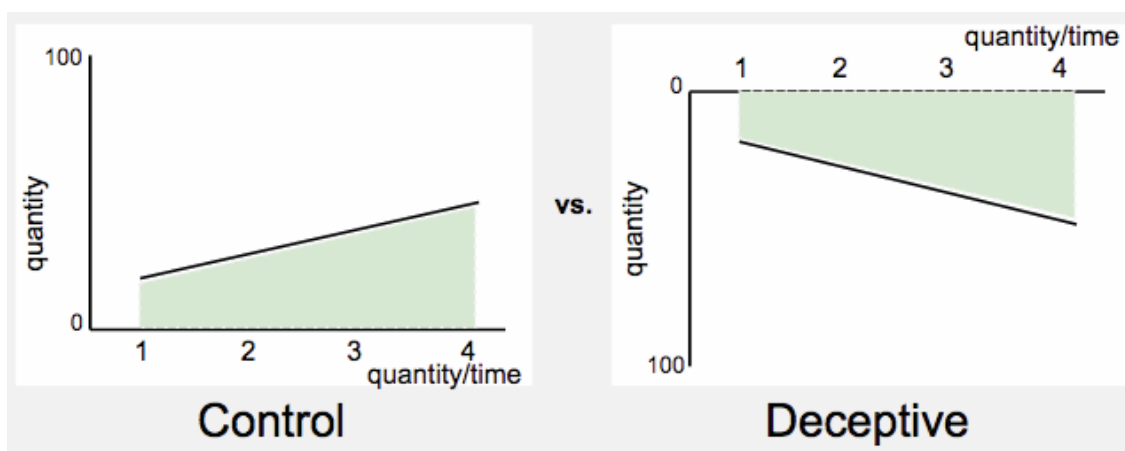


Figure 5. Illustration showing Inverted Axis distortion, which leads to message reversal type of deception.

**Detecting Deception**

- メッセージレベルの歪みを検出する質問を以下のように分類
  - ✓ **Message Exaggeration/Understatement**
    - この種の歪みは、メッセージの事実は歪められず、誇張もしくは過小評価される
    - 本研究では、「How much (どれぐらい大きいか)」形式の質問を 5 段階で評価させ、誇張/過小評価を検出する
    - 誇張/過小評価の程度を検討するために、正しく表現されたグラフと誤って表

現されたグラフを用い、被験者間で比較した

✓ **Message Reversal**

- この種の歪みは、メッセージの事実自体を歪める
- 本研究では、「What（何が読み取れるか）」形式の質問を設定
  - ◇ メッセージの解釈として、「正解、（歪みから導かれる）不正解、どちらでもない」の3つの選択肢から選択
- 被験者間の回答の正確さを比較することで、歪められた視覚化の影響があるかどうかを検討する

**Measure of Individual Differences**

- 視覚化が人の知覚に与える影響を検討する上で重要な点の一つは、提示された情報を読み取り、処理して、根本的なメッセージの解釈をする能力を理解すること
  - 本研究では、参加者の「図表への親密度」と「情報処理能力」の2段階によって、この点を追加する
- 個人差を判断するもう一つの変数は「認知の必要性」
  - 認知の必要性：社会心理学で用いられる、個人が努力的認知活動に傾く程度を特徴づける人格特性
  - 本研究では、18項目で構成される、最も一般的な「認知の必要性」尺度の一つ（Petty & Cacioppo, 1986）を使用
- これらとは別に、年齢、性別、学歴などの人口統計情報を記述させた
  - 高等教育は高いビジュアルリテラシーと、グラフや統計による推論能力につながる可能性があるため、主に学歴を分析した

**User Study**

- 2つの予備調査
  - 1つ目：参加者の反応に歪み手法の影響があるか
    - ◇ 現実世界の歪められた視覚化（Figure 1-a）を用い、提示された情報の誤った解釈を検出する質問に回答させた
    - ◇ 結果、歪められた視覚化は、読み手の誤った解釈を導く可能性が示された
  - 2つ目：実験的に生成されたシナリオとデータにおいても、同様の歪み手法で同様の影響が見られるか
    - ◇ 1つ目の予備調査と同様に、参加者は誤った解釈を行った
    - ◇ 1つ目と2つ目の予備調査の結果に違いは見られなかった
- 本調査
  - 2種類の歪みの影響をそれぞれ検討する2つの実験
    - ◇ メッセージの誇張／過小評価

- “truncated axis”, “aspect ratio”, “area”のグラフについて、それぞれ一つずつ検討
- ◇ メッセージの反転
  - “inverted axis”のグラフについて検討

### Experiment Setup

- Amazon Mechanical Turk (MT) を使用
  - 全体の参加者は 330 名
  - 各実験は 5~10 分

### Procedure (Figure 6)

- Stage 1 : 同意書に同意した上で、年齢、性別、教育レベルなど基本的な人口統計情報を記入
- Stage 2 : 視覚化能力テストを提示
  - 情報処理課題を含む (Micallef, Dragicevic, & Fekete, 2012; Lengler, 2006)
- Stage 3 : グラフの親密度テストを提示
  - 特定の種類のグラフ (棒グラフ、折れ線グラフ、面グラフ) の例を提示
  - それぞれにおいて「以下の図表にどれくらい精通していますか?」に対し、「全くなじみがない (1)」から「非常になじみがある (5)」の 5 段階で評定
- Stage 4 : 提示グラフ、提示されたグラフに関する簡単な概要、メッセージレベルの質問、アテンションチェックを提示
  - 8 つ (4 つのグラフ×統制 or 操作) のグラフのうち、どれか一つが提示された
  - ✓ メッセージの誇張/過小評価に焦点を当てた実験
    - ◇ 棒グラフ、バブルチャート : 提示された変数の比較
      - 「[文脈]の観点から[変数 A]は[変数 B]と比べてどれくらい良いと思いますか」

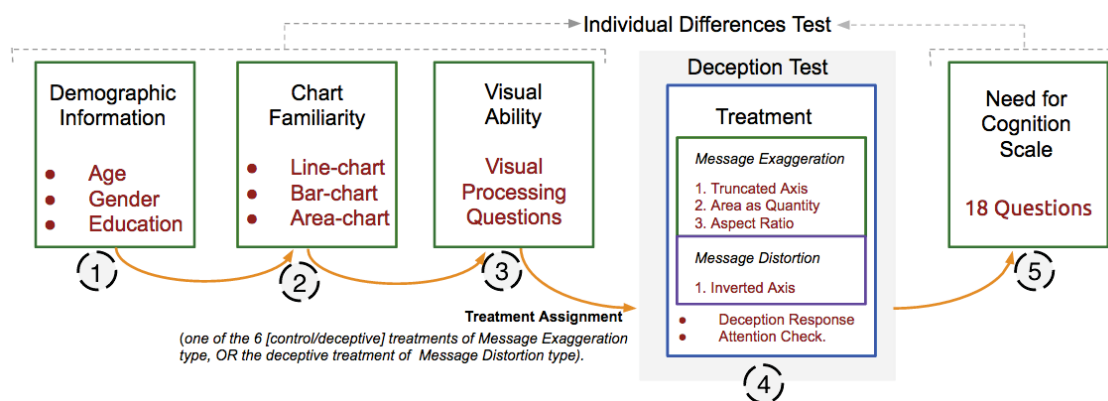


Figure 6. Various stages of the experiment. Stages 1, 2, 3 and 5 correspond to the individual differences test, while 4 corresponds to deception test.

- ◇ 折れ線グラフ：時間の経過によって、変数が改善した程度
  - 「[期間]の間の[文脈]の観点から、[変数 A]はどれくらい改善したと思いますか」
- ✓ メッセージの反転に焦点を当てた実験
- ◇ 折れ線グラフ：
  - 「[変数 A]が[変数 B]に与える影響について、何が言えますか」
- Stage 5：認知の必要性尺度
  - 18 項目、5 段階評定 (Petty & Cacioppo, 1986)

## Results

- 歪められた視覚化に関して、読み手を騙すという観点から、異なる歪み手法の影響力を検討

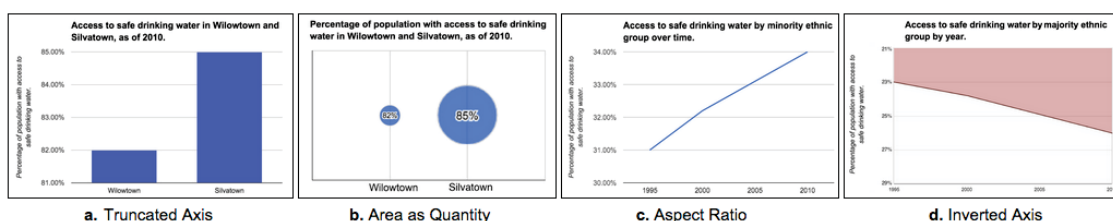


Figure 7. Deceptive visualization examples with corresponding distortion technique used in the study.

Examples (a), (b) and (c) are used for message exaggeration/understatement technique, and (d) for message reversal.

### Message Exaggeration/Understatement

- この種類の歪みの影響を検討するため、縦軸の操作された棒グラフ、面積の違うバブルチャート、アスペクト比の異なる折れ線グラフを使用 (Figure 7-a, b, c)
  - それぞれにおいて、統制条件と操作条件の 2 種類のグラフを作成
- 参加者は 250 名
  - 6 つのグラフ (3 つのグラフ×2 条件) のうち、1 つを提示
  - アテンションチェックに無回答、もしくは正答しなかった参加者を除外
  - 最終的な人数を Table 1 に示す

Treatment	Truncated Axis	Area as Quantity	Aspect Ratio
Control	37	40	38
Deceptive	43	40	42

Table 1. Distribution of participants (*who answered all the attention check questions correctly*) by treatment (Control, Deceptive) and distortion technique (Truncated Axis, Area as Quantity, Aspect Ratio).

### Effect of Distortion on Response

- 全ての種類のグラフで、参加者の回答に歪められたグラフが大きく影響していた



- 歪められたグラフを見た参加者は、根本的なメッセージを誇張して認識し、「How much」形式の質問に対し、より高い評価を付けた (Table 2, Figure 8)

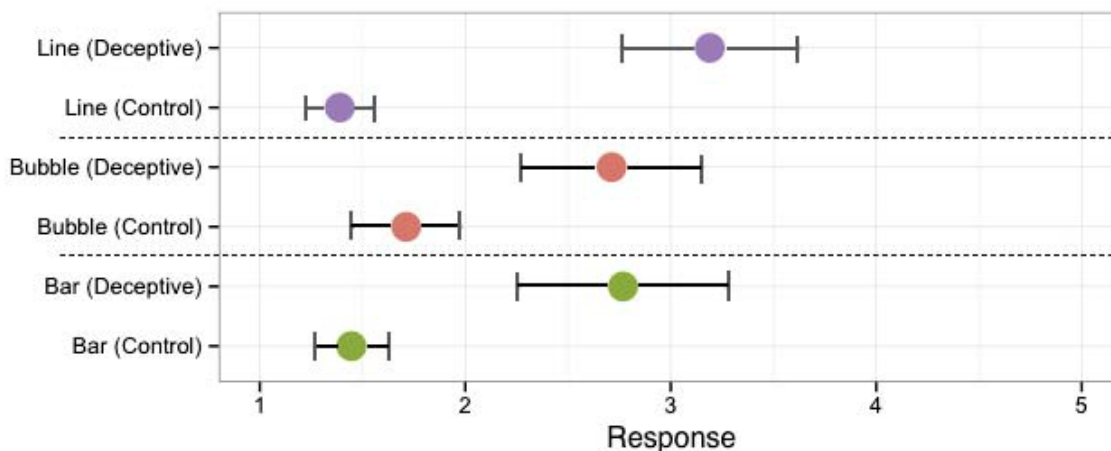


Figure 8. Average participant response with 95% confidence interval, when exposed to a treatment.

Treatment	Distortion Technique	Average Response
Line (Deceptive)	Aspect Ratio	3.19, 95% CI [2.76, 3.62]
Line (Control)	Aspect Ratio	1.39, 95% CI [1.23, 1.55]
Bubble (Deceptive)	Area as Quantity	2.71, 95% CI [2.27, 3.15]
Bubble (Control)	Area as Quantity	1.71, 95% CI [1.45, 1.98]
Bar (Deceptive)	Truncated Axis	2.77, 95% CI [2.26, 3.28]
Bar (Control)	Truncated Axis	1.45, 95% CI [1.27, 1.62]

Table 2. Average participant response (with 95% CI) to the "how much" question, by treatment (Control, Deceptive) and distortion technique (Truncated Axis, Area as Quantity, Aspect Ratio). (*minimum = 1, maximum = 5*)

- 回答の違いを検討するため、各グラフにおいて、マンホイットニーの U 検定を実施
  - 3つのグラフの全てで有意な差が見られた
    - ✓ aspect ratio :  $U = 1409$ ,  $Z = 5.88$ ,  $p < 0.0001$ ,  $r = 0.66$ 
      - Line (Control) の平均ランク : 24.42
      - Line (Deceptive) の平均ランク : 55.04
    - ✓ area as quantity :  $U = 1121$ ,  $Z = 3.08$ ,  $p = 0.0007$ ,  $r = 0.34$ 
      - Bubble (Control) の平均ランク : 32.47
      - Bubble (Deceptive) の平均ランク : 48.52
    - ✓ truncated axis :  $U = 1144$ ,  $Z = 3.36$ ,  $p = 0.0003$ ,  $r = 0.37$ 
      - Bar (Control) の平均ランク : 31.08
      - Bar (Deceptive) の平均ランク : 48.60

### Message Reversal

- この種類の歪みの影響を検討するため、Y軸の反転された折れ線グラフを使用 (Figure 7-d)
  - 統制条件と操作条件の2種類のグラフを作成
- 参加者は80名
  - 2種類のグラフのうち1つを提示
  - ノイズ除去を行い、分析には78名を使用
    - ◇ 統制条件が40名、操作条件が38名

### Effect of Distortion on Response

- Figure 7-dのグラフに基づいて「安全な飲料水へのアクセス状況について、時間が経つにつれて何が言えるか」を質問
  - 正しい解釈 (改善された)、誤った解釈 (減少した)、どちらでもない (分からない) から選択させた
- 各条件の回答者数 (Table 3)

Treatment	Selected	Correct	Incorrect	Uncertain
Control	40	39 (97.50%)	1 (2.5%)	0
Deceptive	38	7 (18.42%)	30 (78.95%)	1 (0.02%)

Table 3. Response distribution (Correct, Incorrect, Uncertain) by treatment type for the *Inverted Axis* distortion.

- 歪められた視覚化による参加者の反応の違いを検討するため、フィッシャーの直接確率検定を拡張した Fisher-Freeman-Halton 検定を実施
  - 結果、統計的に非常に有意であり、参加者の反応に対する歪められた視覚化の影響が示された ( $p < .0001$ )

### Analyzing Individual Differences

- 個人差 (Stage 1, 2, 3, 5) が、歪められた視覚化に影響される程度に関連するかどうかを検討する
  - 個人差は主に行動研究で検討され、個人差が、個人の反応に影響を及ぼすことが示されている (Nov & Arazy, 2013)
- 個人差の独立性を確認するため、要因間の相関を分析
  - それぞれで得られた値を高低で分類
    - ✓ Stage 1 (学歴) :
      - 低学歴 : 初等、中等
      - 高学歴 : 大学、大学院修士、大学院博士
    - ✓ Stage 2 (視覚化能力) :
      - 低視覚化能力 : 正答できた質問が3つ未満

- 高視覚化能力：正答できた質問が3つ以上
- ✓ Stage 3 (グラフに対する親密度) :
  - 親密度低：1~3
  - 親密度高：4~5
- ✓ Stage 5 (認知の必要性； range = 18~90) :
  - 認知の必要性が低い：合計値が54未満の参加者
  - 認知の必要性が高い：合計値が54以上の参加者

- 誇張／過小評価に関する実験で用いた各グラフにおける個人差要因ごとの平均値 (Table 4, Figure 9)

- Table 4 : それぞれの平均値と信頼区間
- Figure 9 : 学歴、視覚化能力、親密度、認知の必要性
  - ◇ データは、操作条件における参加者で、個人差要因低群と高群の平均値
  - ◇ 線それぞれ、線グラフ (青)、バブルチャート (赤)、棒グラフ (黄)

Attribute	Level	Line chart	Bubble chart	Bar chart
Education	Low	3.93, [3.29, 4.56]	3.07, [2.42, 3.72]	2.73, [1.91, 3.54]
	High	2.85, [2.36, 3.33]	2.50, [1.93, 3.05]	2.73, [2.20, 3.24]
Chart Familiarity	Low	2.72, [1.91, 3.53]	2.60, [1.73, 4.49]	4.34, [3.98, 4.69]
	High	3.41, [2.96, 3.89]	2.71, [2.22, 3.21]	2.42, [1.93, 2.94]
Visual Ability	Low	2.88, [1.95, 3.79]	2.72, [1.75, 3.69]	4.07, [3.36, 4.79]
	High	3.33, [2.86, 3.79]	2.68, [2.21, 3.18]	2.23, [1.73, 2.74]
Need for Cognition	Low	3.40, [2.63, 4.20]	3.10, [2.08, 4.15]	3.22, [2.10, 4.32]
	High	3.18, [2.71, 3.68]	2.56, [2.10, 3.03]	2.67, [2.12, 3.20]

Table 4. Average participant response (with 95% CI) across various individual differences attributes, grouped by the deceptive chart type.

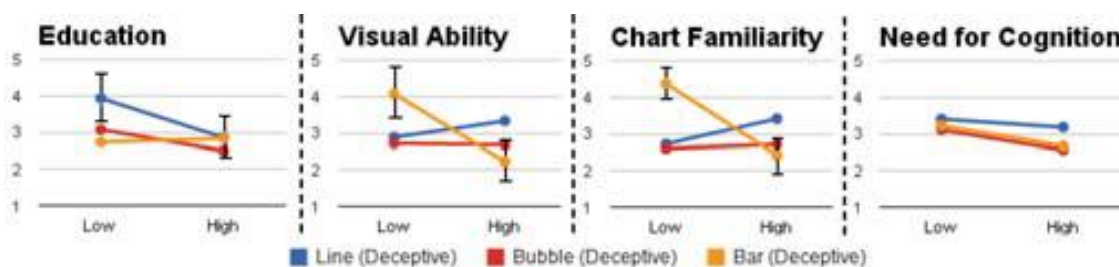


Figure 9. Comparison of average user response across deceptive charts, by the low/high level of individual differences factors.

- 個人差と参加者の回答の間に明確な傾向はない (Figure 9)
  - 唯一見られる主な傾向は、棒グラフが提示されたときには、視覚化能力と親密度が影響すること
  - 信頼区間は重複しているが、折れ線グラフが提示されたときには、学歴が影響する
  - 他のグラフでは信頼区間が広く、重なっており、影響が明確ではないため、グラフ

には示していない

- 視覚化能力とグラフの親密さは、似た結果を示したため、2 要因間の相関分析を行った
  - 結果、低い相関を示した ( $p = .171$ )
- 学歴と認知の必要性は、似たような傾向を示しているが、信頼区間がかなり重複しているため、これらについて強い主張をすることは困難
- 全ての個人差要因間で相関分析を行ったが、どこも強い相関を示さなかった
  - 予測に反して、個人差の分析は、明確な結論に至らなかった

## Discussion

- 本研究の主な目的：
  - データと統計のグラフ表現でよく見られるいくつかの歪み手法によって、人々が騙されるかどうか、また、どの程度騙されるかを検討すること
  - この影響が、個人差によって調整されるかどうかを検討すること
- 結果、これらの表現は、読み手の誤った解釈を引き出し、その影響もかなり大きいことが示された
  - 5 段階評定で 2 つの変数を比較する場合、歪められたグラフは、歪められていないグラフより、58.5%~129.5%大きい反応を導いた
  - メッセージの誇張/過小評価に関する 3 つのグラフの中での影響の大きさ
    - ◇ 折れ線グラフ > 棒グラフ > バブルチャート
  - これらのグラフはより顕著な影響があることが示された
    - ◇ メッセージの反転に関する実験についても同様のことが確認された
      - 誤った回答をした参加者の割合は、操作条件では 97.5%だったが、統制条件では 18.4%だった
- 個人差の分析では、決定的な結果は見られなかった
  - ただ、いくつかの個人差は、特定の種類のグラフに影響しているように見えた
- 歪み手法、図表の種類、個人差の関係については、さらなる研究が必要
  - 解釈に対する影響が、図表の種類に依存するのか、使用されている手法にのみ依存するのか
  - 参加者の一般的なグラフリテラシーが回答と関連するか