The curse of three dimensions: Why your brain is lying to you

SUSAN VANDERPLAS, HEIKE HOFMANN and DIANNE COOK, Iowa State University

One of the basic principles of information graphics is that the data should be accurately reflected in the chart. Tufte's lie factor was created with the idea that graphs that do not represent the underlying data accurately should be avoided. In this paper, we examine a second level of graph distortion that occurs during the perceptual process. The human visual system is largely optimized for perception of three dimensions. Generally, the brain processes potential ambiguities in the rendering as the most-common three-dimensional object. This can lead to visual distortions, such as occur with the Necker figure or in the Müller-Lyer illusion. We discuss the underlying psychological mechanisms for the distortions, examine the effect these distortions have on judgments, and consider the implications for graph design. Using the sine illusion as a case study, we quantify the effects of the distortion that create a "perceptual lie factor" for the sine illusion.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Screen Design*; *User-centered design*; G.3 [**Mathematics of Computing**]: Probability and Statistics—*Statistical computing*; J.2 [**Computer Applications**]: Physical Sciences and Engineering—*Mathematics and statistics*

General Terms: Design, Human factors, Reliability

Additional Key Words and Phrases: Information graphics, optical illusion, perceptual theory, visualization

ACM Reference Format:

VanderPlas, S., Hofmann, H., and Cook, D. 2014. The Curse of Three Dimensions: Why Your Brain is Lying to You. ACM Trans. Appl. Percept. 1, 1, Article 1 (January 2014), 18 pages.
DOI = 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

Author's address: S. VanderPlas, Snedecor Hall, Iowa State University, Ames, IA 50010; email: skoons@iastate.edu

DOI 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

1. INTRODUCTION

Graphics are one of the most powerful tools researchers have to communicate results to wide audiences. They are easier to understand than tables or verbal descriptions[Larkin and Simon 1987], easier to remember than words alone [Mayer and Sims 1994], and provide information that can be perceived and used with minimal additional cognitive load [Zhang and Norman 1994]. Informative graphics differ from visualizations, sketches, and diagrams in that they present visual summaries of data, using summary functions to map the data graphically, preserving the relationship between two variables using spatial information. That is, unlike visualizations, sketches, and diagrams, spatial relationships presented in information graphics are functions of the data, representing numerical quantities (within the limits of image resolution) accurately. Conveying numerical relationships using spatial information makes use of the brain's visual processing ability, freeing working memory to interpret and make connections between graphics and written interpretations.

As information graphics utilize spatial relationships to convey numerical information, it is particularly important that these relationships accurately portray the data [Tufte 1991]. In many cases, attempts to make a chart visually interesting compromise this goal: using three dimensions when only one dimension is required, or adding unnecessary perspective views to a chart can distort the numerical information the plot was designed to convey. In his book, Tufte argues that some graphics (many of which might be better classified as visualizations) do not accurately reflect the data because they sacrifice numerical accuracy for visual appeal. In order to quantify this loss of numerical accuracy, Tufte created the **lie factor**, which compares the effect size shown in the image to the effect size shown in the data, so that a lie factor much greater than 1 indicates a picture that over-emphasizes an effect, and a lie factor much smaller than one indicates a picture which minimizes an effect (values between .95 and 1.05 are typically acceptable). While Tufte's lie factor is an effective measurement of the accuracy of the transition from data to graphics, we must also be concerned about the transition from graphics to the brain, as we are relying on the brain's visual processing ability to efficiently encode numerical relationships.

Ideally, charts not only represent the data accurately, but also allow readers to draw accurate conclusions. Generally, the human visual system is quite good at accurately interpreting charts [Cleveland and McGill 1984; Kosara and Ziemkiewicz 2010], but we need to be aware of contextual misperceptions that lead us to the wrong conclusion. While there are relatively few examples of the effect of optical illusions and other misperceptions on information graphics, Amer [2005] and Poulton [1985] have documented the effect of the Poggendorff illusion on line graphs in different contexts. Tufte's lie factor does not give us any insight about the accuracy of the transition between chart and brain, but a similar factor could be created to measure the accuracy of this transition. In this paper, we examine a situation in which low-level human perceptual processes interfere with making accurate judgments from displays and suggest an experimental methodology for estimating the psychological "lie factor" of a chart due to a specific conceptual misperception: the sine illusion.

1.1 The Curse of Three Dimensions

The human visual system is largely optimized for perception of three dimensions. Biologically, binocular vision ensures that we have the necessary information to construct a functional mental representation of the three-dimensional world, but even in the absence of binocular information the brain uses numerous heuristics to parse otherwise ambiguous two-dimensional retinal images into meaningful three-dimensional information. Predictably, however, these heuristics are not without drawbacks; the same two-dimensional neural representation might correspond to multiple three-dimensional objects, as in the well-known Necker Cube (Gregory 1968; shown in figure 1). Additionally, the same three-dimensional object often has infinitely many two-dimensional representations, for instance, when viewed from different angles. Many optical illusions occur due to the transition from a two to three dimensions (or from three dimensions to two dimensions)[Gregory 1968]. The necker cube has a single two-dimensional representation corresponding to two three-dimensional objects which are both equally salient. As a result, the brain does not prefer one interpretation over the other and instead continuously switches between interpretations. Impossible objects, such as the

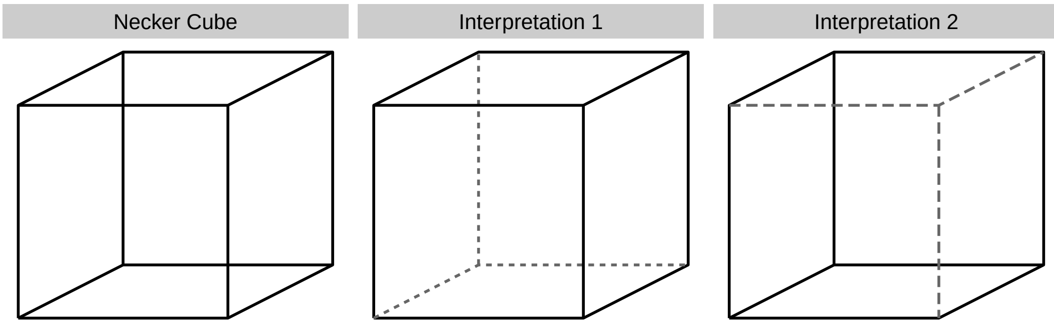


Fig. 1. The Necker Cube is a so-called “ambiguous object” because two different transparent objects produce the same retinal image (and thus the same perceptual experience). Commonly, the image seems to transition instantaneously from one possible mental representation to the other.

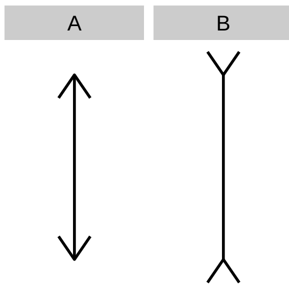


Fig. 2. The Müller-Lyer illusion. The central segment of figure A is perceived as shorter than the central segment of figure B, even though the two are actually the same length.

Penrose triangle [Penrose and Penrose 1958], are two-dimensional images of objects that appear to be impossible in three dimensions (sometimes, these objects can be represented in three dimensions, but only appear to be impossible from a certain angle). Impossible objects produce a conflict between the brain’s three-dimensional representation of a two-dimensional figure and the brain’s experience with the physical world. This conflict between the constraints of physical reality and a depicted image can be quite compelling and is an important component in the work of artists like M.C. Escher [Seckel 2007].

In non-illusory contexts, experience with the real world informs the choice between multiple possibilities of rendering a three-dimensional object from the same two-dimensional representation. This indicates that processing occurs “top down” in that our previous experience influences our current perceptions. Without this top-down influence, the brain would not be able to map a two-dimensional image back to a three-dimensional object. One of the most well studied examples of the influence of top-down processing is the Müller-Lyer illusion, shown in figure 2.

In the Müller-Lyer illusion, two vertical line segments are shown with arrows extending from both ends; one segment forms an acute angle with the arrows, the other segment forms obtuse angles with the arrows. The line segment adorned with arrows that form an acute angle appears to be shorter than the linen segment which forms an obtuse angle with the arrow segments.

One explanation for the Müller-Lyer illusion [Gregory 1968] is that the brain interprets the ambiguous lines as a common three-dimensional object common to everyday experience: corners of a room. Figure 2A occurs when viewing the outside corner of a rectangular prism, figure 2B occurs when viewing the prism from the inside. In regions which do not commonly have rectangular buildings, the illusion is significantly less pervasive [Ahluwalia 1978]. Figure 3 provides one possible context that would lead to the Müller-Lyer effect. This real-world experience carries with it

Fig. 3. Real-world context that gives rise to the Müller-Lyer illusion. The highlighted areas correspond to the parts of the Muller Lyer illusion, and while the two arrows are obviously the same size in the real world, the black arrow takes up much more visual space than the white arrow.

an inferred perspective - when the arrows point inward, the object is typically closer than when the arrows point outward, which causes the brain to interpret the outward-pointing figure as larger when the retinal size of the two objects is identical. The perspective cues which contribute to the Müller-Lyer illusion allow for an accurate neural representation of the object in context; when misapplied to two-dimensional stimuli, these cues are responsible for the illusion's effect. This inferred "depth cue" [Gregory 1968] is reasonably consistent across individuals, suggesting that the phenomenon has a neurological basis.

A similar effect can also be found in the Necker Cube - whichever face appears to be furthest away also seems larger, even though any two parallel faces are equally sized in the image. This approach has proved to be very advantageous for real world scenarios [Gregory 1968], as pictures of real objects are seldom ambiguous. This strategy also allows for high performance with limited neural bandwidth.

1.2 Three Dimensional Context of the Sine Illusion

While the classic Müller-Lyer illusion is seldom a factor in information charts, there are other illusions caused by the interpretation of a two-dimensional stimulus in the context of three-dimensional objects, leading to a distortion in the mental representation of the original stimulus. The sine illusion (also known as the line width illusion: VanderPlas and Hofmann 2014; Day and Stecher 1991) is one example of this phenomenon which occurs frequently in information graphics.

Figure 4 shows the sine illusion in its original form [Day and Stecher 1991] as straight vertical lines of the same length with a sinusoidal mean function. In this illusion, the vertical lines in the center of the figure appear much shorter than the vertical lines at the peak and trough of the sine curve. The illusion still persists when the image is rotated by 90°. Even when viewers are aware of the illusion's presence, it is close to impossible to overcome mentally.

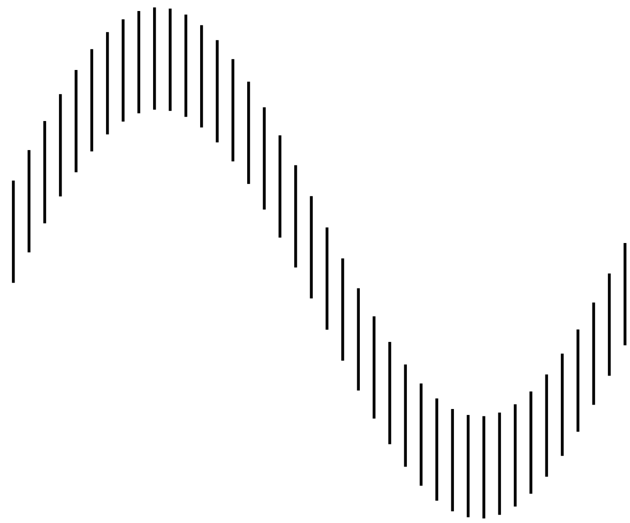


Fig. 4. The classic sine illusion. Each vertical line has the same length, though the lines at the peak and trough of the curve appear longer.

The problem that the sine illusion presents in information graphics is well documented [Cleveland and McGill 1984; Robbins 2005]. One such example is the "Balance of Trade" from Playfair's Statistical Atlas [Playfair 1786], as shown in figure 5. The balance of trade in 1765 seems to be approximately the same as the balance of trade in the years immediately preceding that year; this is in fact extremely misleading (using a straightedge along the chart vertically will demonstrate the issue).

In both figures, the illusion appears when the vertical length displayed in the chart does not match the perceived information. Like the Müller-Lyer illusion, the illusion is pervasive and very difficult to "un-see" or mentally correct. The sine illusion, which is also known as the line-width illusion, has also been documented in parallel sets plots [Schonlau 2003; Hofmann and Vendettuoli 2013] and occurs when there is a nonlinear function with a large change in absolute slope; this change in slope can mask or exaggerate changes in variance. The illusion is also affected by the aspect ratio of the image and the aspect ratio of the chart's coordinate system. An interactive demonstration of the illusion is available at `http://bit.ly/1ldgujL`; manipulating the length of the lines and the amplitude of the underlying sine function also changes the chart's aspect ratio and the perceived strength of the illusion.

The illusion is not dependent on specifically identifying the vertical distance along a line. Figure 6(a) shows a scatterplot of data with a trend. A loess smooth is used to estimate the trendline. A visual assessment of variability along this trendline might result in a description such as 'homogeneous variance or slightly increasing variance for negative x, followed by a dramatic decrease in vertical variability for positive x'. Once the residuals are separated from the trendline as shown on the right hand side of the figure, it becomes apparent that this first assessment of conditional variability was not correct, and the steady decrease along the horizontal axis becomes visible.

Cleveland and McGill [1984] determined that comparison of the vertical distance between two curves is often inaccurate, as "the brain wants to judge minimum distance between the curves in different regions, and not vertical distance". While they do not explain a reason for this tendency, introspection does support their explanation: we judge the distance between two curves based on the shortest distance between them, which geometrically is the distance along the line perpendicular to the tangent line of the curve. This comparison holds with scatterplots (such as figure 6) because when the points are dense, we examine variability by looking at the upper and lower contours of the data.

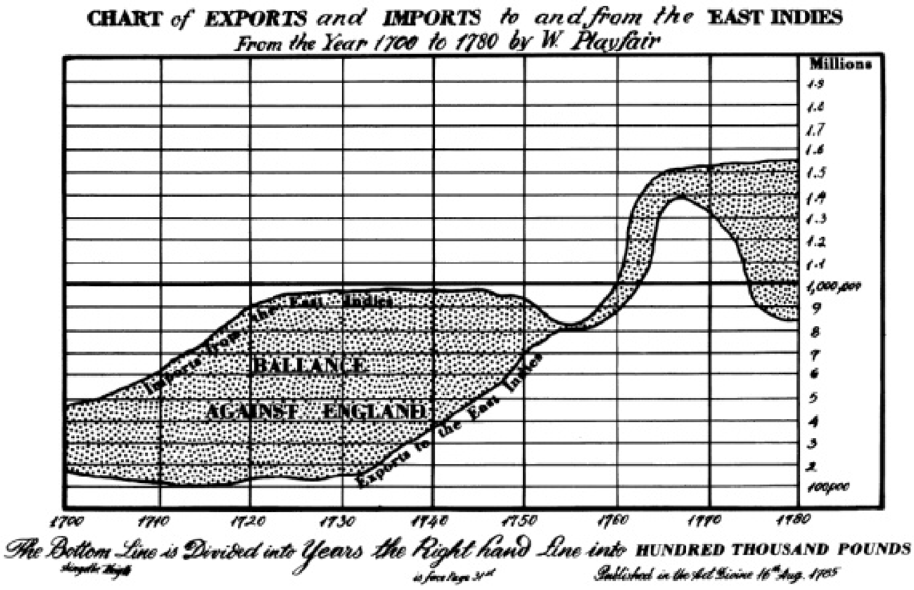
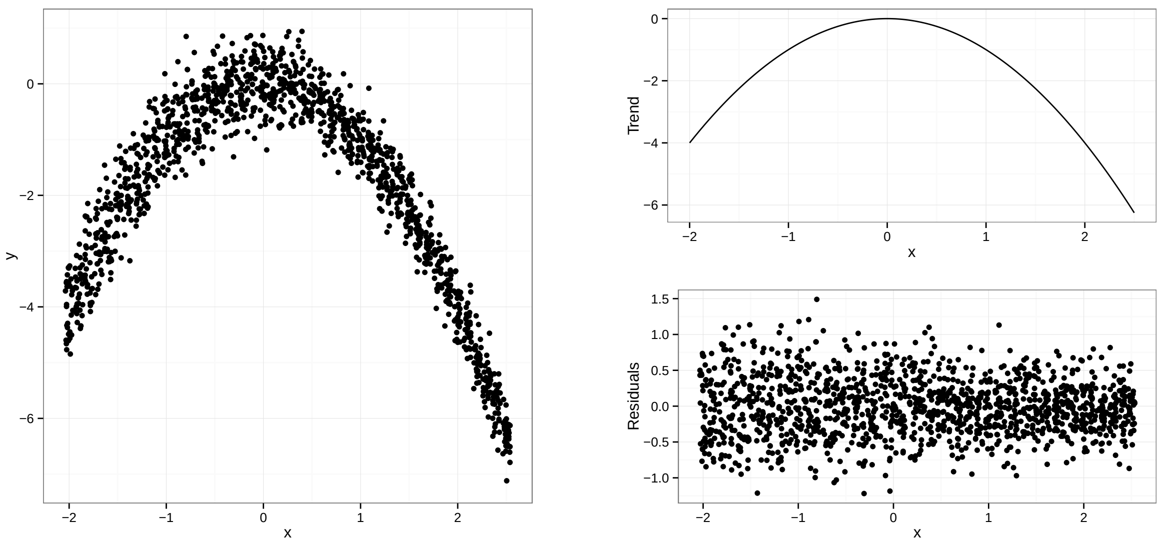


Fig. 5. Playfair's chart of trade between the East Indies and England, 1700-1780. The trade balance is influenced by the sine illusion: the difference between imports and exports in 1763 does not appear to be the same size as that in 1745, though the vertical distance is approximately the same.



(a) Data

(b) Trend and Residuals

Fig. 6. Describe the conditional variability of the points along the *x* axis in (a). Is your description consistent with the residual plot in (b)?

Day & Stecher [1991] suggest that the sine illusion is similar in principle to the Müller-Lyer illusion, attributing it to the perceptual compromise between the vertical extent and the overall dimensions of the figure. The sine illusion is similar to the Müller-Lyer illusion in another way, as well – there are three-dimensional analogues of the two-

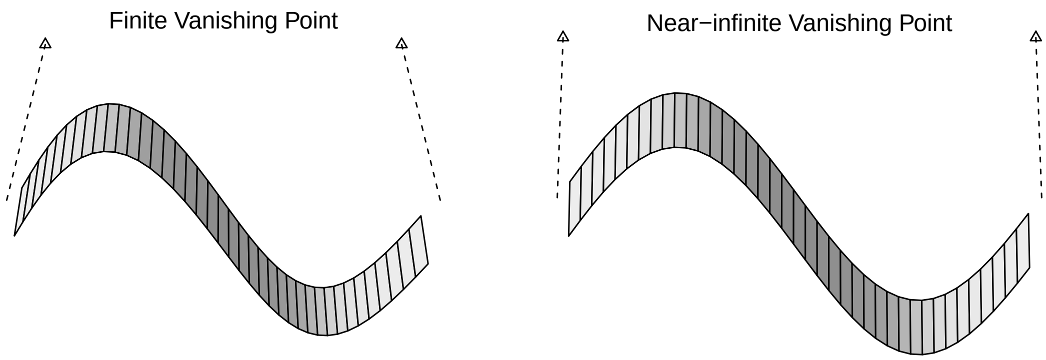


Fig. 7. Three dimensional context for the sine illusion. The second figure has a vanishing point closer to infinity, and very closely resembles the form of the classic sine illusion.

dimensional image that may influence the perceptual context. One of these contexts is shown in figure 7, generated from the same function shown in the two-dimensional analogue, figure 4, but with the length projected onto a third dimension. While the images do not match exactly, the similarities are striking. Additionally, the tendency to judge vertical distance using the extant width noted in Cleveland and McGill [1985] corresponds to the measurement of depth in the three-dimensional image. The main difference between the first three dimensional image shown in figure 7 and the original image is that the lines connecting the top and bottom sections of the curve are slightly angled in the three-dimensional version; this is due to the perspective projection used to create the image and the corresponding angles of rotation chosen such that the entire surface is visible.

As the vanishing point moves further away from the viewer and the 3d projection decreases in strength, the three-dimensional reconstruction of the image converges to figure 4. The second image in figure 7 shows a weaker 3-dimensional projection that is much closer to figure 4, however, the three-dimensional contextual information provided by the shading removes much of the illusion's distortions. This is similar to the Müller-Lyer illusion, as figure 3 is not at all ambiguous because the contextual depth information provided by the rest of the surface of the house is sufficient to remove the illusion that the closer corner is in fact larger due to the perspective.

1.3 Case Study: Three Dimensions and the Sine Illusion

Further evidence that places the sine illusion firmly into the area of a 3d contextual illusion is given by the non-response to the illusion by individuals with depth-deficient vision. While it is difficult to provide experimental evidence suggesting that the sine illusion is due to depth perception directly, it is possible to examine whether the illusion occurs in people who do not have binocular depth perception. Conditions such as amblyopia (lazy eye) and strabismus (crossed eyes), when not corrected within a critical period during development [Hubel and Wiesel 1970], can result in weakened or absent depth perception [Henson and Williams 1980; Parker 2007; Holopigian et al. 1986]. In many cases, use of partial patching and early surgery can correct these problems before the critical period lapses, but before protocols were well established, this was not always completely successful.

We examined the effect of the sine illusion on DW, who has minimal depth perception due to strabismic amblyopia. DW was diagnosed as a young child, and prescribed complete patching to strengthen her initially non-dominant eye. As a result of the patching, DW developed near-independent control over both eyes (doctors now recommend partial patching as a result of this problem). She has 20/20 vision, and can wear glasses to correct the strabismus, but generally does not because they are not necessary for her to see well. DW is right-eye dominant in most contexts, but is left-eye dominant for driving, and can switch which eye is in focus at will.

We asked DW to view a subset of the sine illusion stimuli used in the experiment described in the next section, as well as the Müller-Lyer illusion, identifying the illusions as having lines that appeared the same length or different lengths (the stimuli are included in the appendix). DW identified both uncorrected sine illusion graphs as having lines of the same length, indicating that she did not appear susceptible to the sine illusion. In addition, DW identified the partially corrected images as having the same line length, indicating that the corrected image would produce similar conclusions as the uncorrected image (in this, she was not different from those with normal binocular vision). In fact, DW only identified the fully corrected $y = \exp(x)$ image as having lines of different length.

In addition to DW's resistance to the effects of the sine illusion, she also was not fooled by the Müller-Lyer illusion, instantaneously identifying the lines as the same length. This suggests that these two illusions are related to the presence of binocular depth perception, perhaps mitigated by experience.

One difference between the sine illusion and the Müller-Lyer illusion that may influence the tendency to see a three-dimensional "ribbon" instead of the two-dimensional sine curve is that the vertical lines in the sine illusion are ambiguously oriented - there is an entire plane of possible three-dimensional reconstructions for each line, and each possible rotation leads to a line of different length. It is this facet of the image that we believe partially contributes to the ambiguity of the image, though it is not a necessary feature for the illusion to persist, as the illusion also can be found in scatterplots and in "ribbon plots" such as figure 5.

2. MEASURING THE PSYCHOLOGICAL LIE FACTOR EXPERIMENTALLY

2.1 The Psychological Lie Factor

The psychological mechanisms which force three-dimensional context onto two-dimensional stimuli are useful adaptations to a three-dimensional world [Gregory 1968], but they do have disadvantages when applied to abstract two-dimensional stimuli, such as information charts. In order to assess the distortions due to the illusion, we need to quantify this distortion. For comparison, we will work from Tufte's lie factor [Tufte 1991, pg 57], which compares the size of an effect in the data with the size shown in a graphic, and is defined in equation 1.

$$\text{Lie Factor} = (\text{size of effect shown in chart})/(\text{size of effect shown in data}) \tag{1}$$

We will similarly define the psychological lie factor for this illusion as shown in equation 2.

$$\text{Psychological Lie Factor} = (\text{size of effect perceived})/(\text{size of the effect shown in the chart}) \tag{2}$$

A correction factor which focuses on correcting the psychological distortion caused by the sine illusion is detailed in VanderPlas and Hofmann [2014]. This correction factor c_ℓ (the y correction in the aforementioned paper) is applied to the line segments and extends these segments vertically at location x according to equation (3), where w represents a weight factor to allow variation of the strength of the correction. For a differentiable function $f(x)$ (obtained by regression or loess smoothing) with derivative $f'(x)$, the correction is:

$$c_\ell(w) = (1-w) + w\sqrt{1+f'(x)^2} \tag{3}$$

A value of $w = 0$ indicates that there is no correction, and a value of $w = 1$ indicates that the graph is fully corrected. Extending this approach, we can over-correct or under-correct the graph, to test whether the geometric derivation of the correction is sufficient to remove the illusion. The lie factor can then be determined by varying w, so that the lie factor of the plot selected so that the lines appear to be "even" indicates the level of psychological distortion (since lines which are in fact not even but appear to be even would indicate that some distortion occurred within the brain). An example of the correction's effect and various weight factors can be seen in figures 9 and 10.

In order to estimate the psychological lie factor that occurs due to this illusion, we assessed the strength of the illusion experimentally.

2.2 Study Design

The study was designed as a factorial exploration of the factors that contribute to the psychological distortion. We varied the underlying mean function of the stimuli, as well as the strength of the correction described in equation 3.

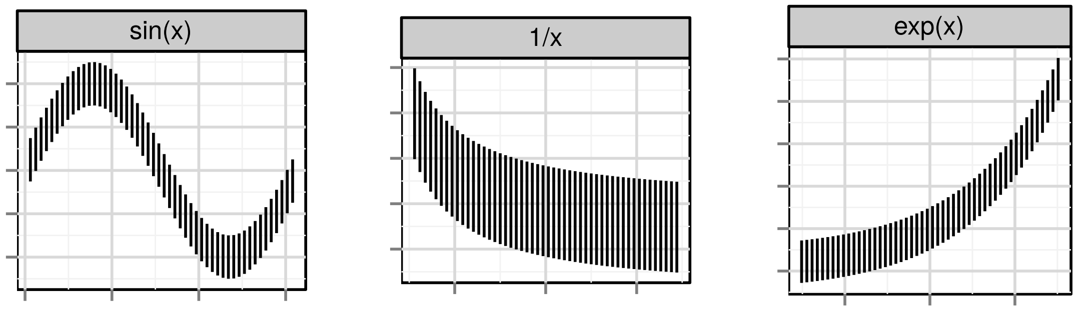


Fig. 8. Mean functions used during the experiment: $\sin(x)$, $\exp(x)$, and $1/x$. These functions are nonlinear, easily differentiable (for the correction factor), and are similar to trends commonly found in information graphics.

Three underlying mean functions were used: $y = 2\sin(x)$, $y = \exp(x/2)$, and $y = 5/(6x)$; varying these functions allowed us to consider whether the psychological lie factor was influenced by the underlying function. The mean functions, shown in figure 8, were chosen from nonlinear functions that occur relatively frequently in statistics and are also easily differentiable (for the correction in equation 3). The scales are chosen, such that the aspect ratio are similar for each set of plots (between 0.75 and 0.85). As no x or y units were provided on the graph, these functional modifications served as experimental controls but did not change the information provided to the participants.

In addition to varying the underlying mean function, we also varied the strength of the correction factor (as described by the parameter w in equation 3). Experimental stimuli consisting of sets of six sub-plots were constructed such that each sub-plot was generated using a different w value between 0 and 1.4. Two of the stimuli used in the experiment are shown in figures 9 and 11. Figure 10 shows the amount of line correction used in each of the sub-plots in figure 9, and the (ordered) w values and corresponding lie factors are shown in table I (row #4).

For each of the stimuli, participants were asked to answer the question: “In which graph is the size of the curve most consistent?”. The phrasing ‘size of the curve’ was chosen deliberately so as not to bias participants to explicitly measure line lengths.

Figure 11 shows another set of these stimuli using a different underlying mean function with the same underlying weight values. As the slope of the mean function has changed, the illusion does appear to be slightly less misleading. Table III (plot #4) shows the weight values (ordered from least to greatest) and corresponding lie factors for this plot; they are lower than the lie factors corresponding to the sine illusion plots shown in figure 9 even though the weight values are the same. The illusion is still present even though the mean function has changed. Our goal is to determine whether the psychological distortion is similar despite the difference in the underlying function.

2.3 Quantifying the Psychological Lie Factor

I think that I have finally figured out what my problem with D_k is – the psychological lie factor is defined as the ratio of the size of the perceived effect and the size of the shown effect. That means that we cannot measure the psychological lie factor without input from somebody who perceives effects. Therefore we should not call D_k the psychological lie factor of plot k. These values are all candidates until somebody picks one of the sub-plots. I think, it’s just a naming convention.

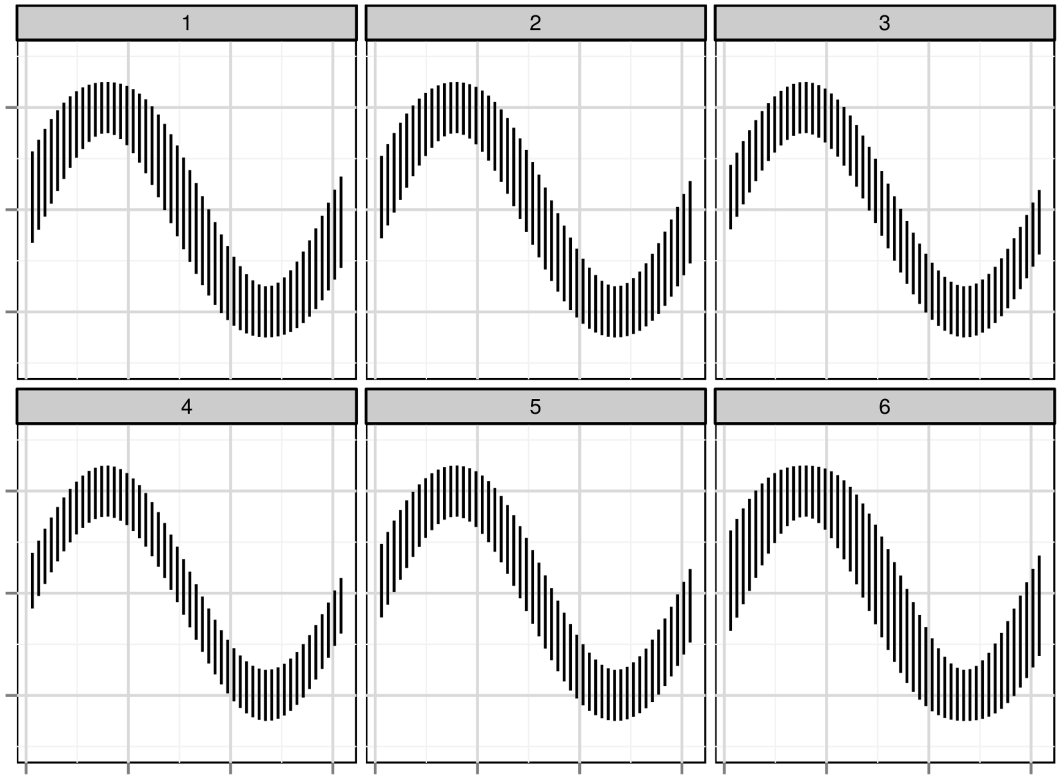


Fig. 9. One of the charts presented to participants through Amazon Mechanical Turk. Figure 10 shows the actual differences in line lengths. This chart corresponds to set #4 in table I. The plots are shown in random order; plot #1 corresponds to $w = 0.9$, plot #2 to $w = 0.7$, plot #3 to $w = 0.3$, plot #4 to $w = 0.1$, plot #5 to $w = 0.5$, and plot #6 to $w = 1.1$.

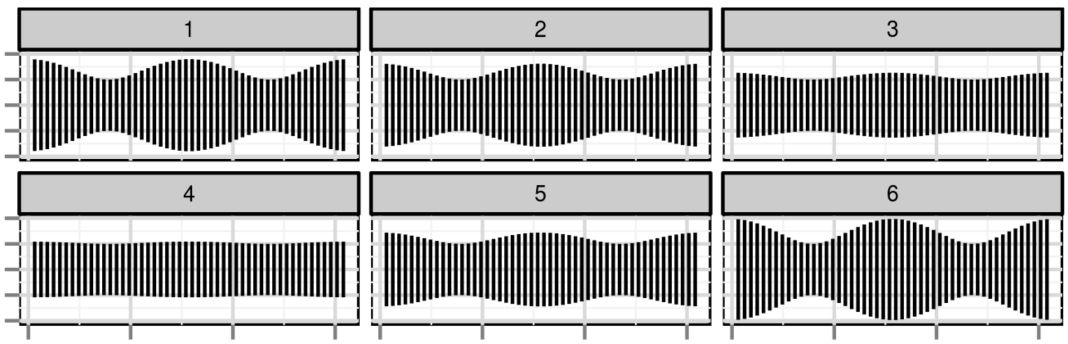


Fig. 10. De-trended line lengths for figure 9, demonstrating the distortion present due to the correction factor in each sub-plot. Comparing the distortion in the chosen sub-plot to the undistorted data produces an estimate of the psychological lie factor.

In an attempt to estimate the psychological lie factor for each of the three functions, we first define the *distortion* D_k for each sub plot $k = 1, \ldots, 6$, as the ratio of the maximum line length to the minimum line length shown in the plot. This results in a positive measure D, for which $D = 1$ indicates no distortion, while a value of D far from one indicates a significant distortion.

$$D_k = (\text{maximum line length in sub-plot } k)/(\text{minimum line length in sub-plot } k) \tag{4}$$

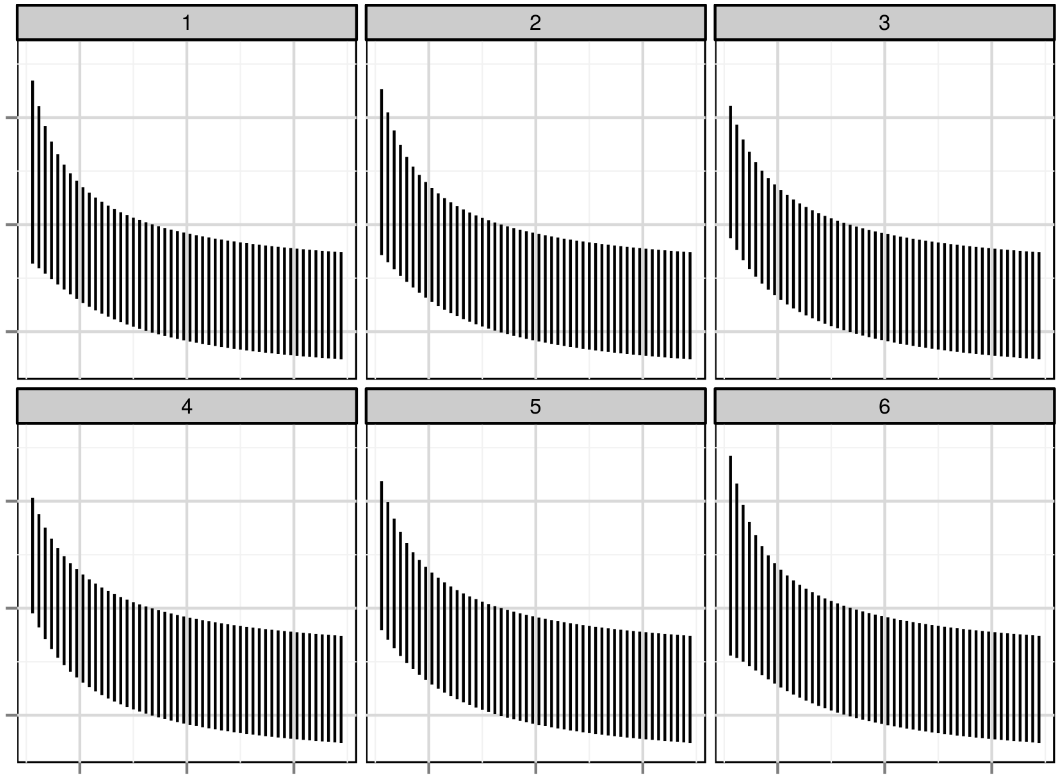


Fig. 11. Another chart presented to participants through Amazon Mechanical Turk. This chart corresponds to set #4 in table III. The plots are shown in random order; plot #1 corresponds to $w = 0.9$, plot #2 to $w = 0.7$, plot #3 to $w = 0.3$, plot #4 to $w = 0.1$, plot #5 to $w = 0.5$, and plot #6 to $w = 1.1$.

In the experiment, we ask onlookers to identify the subplot from the chart with the most consistent lines along the function $f(x)$. This, in turn, identifies the corresponding D_k value as the one which leads to the least amount of perceived distortion, which allows us to estimate the psychological lie factor underlying function $f(x)$. Note that D_k and the weight used for the correction factor w depend on each other (see equation (3)).

what are the results from the experiment for figures 9 and 11? Could we talk through those results to explain our procedure?

One of the stimuli shown to participants is provided in figure 9. A participant might select plot 2 as appearing the most even; we compute the distortion factor $D_2 = 1.80/1.00 = 1.80$, that is, this participant identified a figure with lines varying in length from 1 to 1.8 as appearing to be the most consistent (weight values and the corresponding distortion factors are shown in Tables I, II, and III for sine, exponential, and inverse stimuli respectively.

Based on an internal pilot study, we expected values around $w = 0.8$ to be sufficient to break the illusion, but did not know whether this would generalize to those outside of the information graphics community. In order to pinpoint the weight value necessary to correct the illusion more precisely, we chose twelve sets of six weight values each that were used to produce test plots similar to that shown in figure 9. These sets of weight values were chosen to allow greater precision estimates closer to $w = 0.8$, while still covering the range of w between 0 and 1.4. The sets of w

used are shown in table I, along with corresponding distortion factors D_k for stimuli with underlying function $sin(x)$ (other functions used will have different D_k due to the nature of the correction factor). Plots with weight values spread over the full range of w tested were considered easy "test" plots that could be used for verification purposes, while plots with weight values concentrated near $w = 0.8$ were considered to have higher difficulty (because the sub-plots were very similar). This allowed us to estimate w, and thus D, with higher precision while still exploring the entire parameter space.

Table I. Ordered weight factors and corresponding distortion factors for the sine curve stimuli sets, as computed using equation 4.

set	difficulty	Weight (w) sub-plot 1	2	3	4	5	6	Distortion (D) for sin(x) plots sub-plot 1	2	3	4	5	6
1	test	0.00	0.20	0.40	0.80	1.25	1.40	1.00	1.18	1.35	1.71	2.11	2.24
2	test	0.00	0.15	0.35	0.80	1.20	1.40	1.00	1.13	1.31	1.71	2.06	2.24
3	1	0.00	0.20	0.40	0.60	0.80	1.00	1.00	1.18	1.35	1.53	1.71	1.88
4	1	0.10	0.30	0.50	0.70	0.90	1.10	1.09	1.27	1.44	1.62	1.80	1.97
5	2	0.05	0.30	0.50	0.65	0.80	1.00	1.04	1.27	1.44	1.57	1.71	1.88
6	2	0.10	0.30	0.55	0.70	0.85	1.00	1.09	1.27	1.49	1.62	1.75	1.88
7	3	0.40	0.60	0.70	0.80	0.90	1.05	1.35	1.53	1.62	1.71	1.80	1.93
8	3	0.35	0.65	0.75	0.85	0.95	1.05	1.31	1.57	1.66	1.75	1.84	1.93
9	4	0.35	0.50	0.60	0.70	0.80	0.95	1.31	1.44	1.53	1.62	1.71	1.84
10	4	0.40	0.55	0.65	0.75	0.85	1.00	1.35	1.49	1.57	1.66	1.75	1.88
11	5	0.50	0.65	0.75	0.80	0.90	1.00	1.44	1.57	1.66	1.71	1.80	1.88
12	5	0.50	0.60	0.70	0.75	0.85	1.00	1.44	1.53	1.62	1.66	1.75	1.88

Table II. Ordered weight factors and corresponding distortion for the exponential curve stimuli sets, as computed using equation 4.

set	difficulty	Weight (w) sub-plot 1	2	3	4	5	6	Distortion (D) for exp(x) plots sub-plot 1	2	3	4	5	6
3	1	0.00	0.20	0.40	0.60	0.80	1.00	1.00	1.21	1.42	1.63	1.84	2.05
4	1	0.10	0.30	0.50	0.70	0.90	1.10	1.11	1.32	1.53	1.74	1.95	2.16
5	2	0.05	0.30	0.50	0.65	0.80	1.00	1.05	1.32	1.53	1.69	1.84	2.05
6	2	0.10	0.30	0.55	0.70	0.85	1.00	1.11	1.32	1.58	1.74	1.90	2.05
7	3	0.40	0.60	0.70	0.80	0.90	1.05	1.42	1.63	1.74	1.84	1.95	2.10
8	3	0.35	0.65	0.75	0.85	0.95	1.05	1.37	1.69	1.79	1.90	2.00	2.10
9	4	0.35	0.50	0.60	0.70	0.80	0.95	1.37	1.53	1.63	1.74	1.84	2.00
10	4	0.40	0.55	0.65	0.75	0.85	1.00	1.42	1.58	1.69	1.79	1.90	2.05
11	5	0.50	0.65	0.75	0.80	0.90	1.00	1.53	1.69	1.79	1.84	1.95	2.05
12	5	0.50	0.60	0.70	0.75	0.85	1.00	1.53	1.63	1.74	1.79	1.90	2.05

Each participant was presented with eleven sets of graphs (each "set" made up of six separate plots), consisting of one "easy" test set, five stimuli sets of difficulty level 1 through 5 with the sine curve as the underlying function, and another five graph sets (also of difficulty levels 1 to 5) with either the exponential or the inverse curve as the underlying function.

Difficulty level is determined by the range of weight values shown in the stimuli: easy sets contain plots which span a wide range of weight values, while harder sets span a more narrow range of weight values.

How do you determine difficulty level? You have talked about easier and harder plots, but you have not come up with the level yet. Just define it.

Table III. Ordered weight factors and corresponding distortion for the inverse curve stimuli sets, as computed using equation 4.

set	difficulty	Weight (w)						Distortion (D) for $1/x$ plots					
		sub-plot						sub-plot					
		1	2	3	4	5	6	1	2	3	4	5	6
3	1	0.00	0.20	0.40	0.60	0.80	1.00	1.00	1.14	1.28	1.43	1.57	1.71
4	1	0.10	0.30	0.50	0.70	0.90	1.10	1.07	1.21	1.36	1.50	1.64	1.78
5	2	0.05	0.30	0.50	0.65	0.80	1.00	1.04	1.21	1.36	1.46	1.57	1.71
6	2	0.10	0.30	0.55	0.70	0.85	1.00	1.07	1.21	1.39	1.50	1.60	1.71
7	3	0.40	0.60	0.70	0.80	0.90	1.05	1.28	1.43	1.50	1.57	1.64	1.75
8	3	0.35	0.65	0.75	0.85	0.95	1.05	1.25	1.46	1.53	1.60	1.67	1.75
9	4	0.35	0.50	0.60	0.70	0.80	0.95	1.25	1.36	1.43	1.50	1.57	1.67
10	4	0.40	0.55	0.65	0.75	0.85	1.00	1.28	1.39	1.46	1.53	1.60	1.71
11	5	0.50	0.65	0.75	0.80	0.90	1.00	1.36	1.46	1.53	1.57	1.64	1.71
12	5	0.50	0.60	0.70	0.75	0.85	1.00	1.36	1.43	1.50	1.53	1.60	1.71

After first presenting an easy introductory chart, the order of the plots was randomized across difficulty level as well as function type. The test chart consisted of a set of six sine curves with a very low level difficulty level, and was used as an introduction to the testing procedure. Participants were asked to select a single plot out of the six plots presented as having lines which were the most "consistent".

Is it "even" or "consistent"? What question did you use, exactly? We need to stick to that.
From my emails with Mahbub:
In what picture is the size of the curve most consistent?

2.4 Data Collection

Participants for the study were recruited through the Amazon Mechanical Turk web service, which connects workers with tasks that are not easily automated. In exchange for completing at least 11 trials, participants were paid \$1. Given the anonymity of web-based data collection, we informed participants that a unique IP address was required to participate in the experiment; responses from duplicate IP addresses with different Turk IDs were grouped and only the first set of responses was paid. This procedure was used to deter participants from completing the task multiple times, in order to preserve independence between responses from participants.

After removing data from participants who did not complete at least 10 trials, 106 participants completed 1542 trials.

Due to the experimental design, all participants completed trials with underlying function $y = \sin(x)$, for a total of 815 trials. As each participant who completed only the required 11 trials saw charts with either $y = 1/x$ or $y = \exp(x)$ as the underlying function, each of these functions had fewer trials; 316 and 411, respectively.

I'm just checking on the demographics - we usually collect them in a separate table.

2.5 Analysis

Psychological "Lie Factor". We quantify the amount of distortion across the horizontal range of the curve as the ratio of the maximum line length to the minimum line length for each sub-plot k: $D_k = l_{\max}/l_{\min}$ (see definition (4)).

Let D_{ijk} denote the distortion factor corresponding to participant i's choice of sub-plot k in stimulus set j during a single trial. In this experiment, there are 32 stimuli sets (2 test sets and 10 sets for each of 3 underlying functions), so $1 \leq j \leq 32$.

The correction for the sine illusion we use here extends the line segments, so that for initial line segments of length 1, the correction produces line segments of length greater than or equal to 1. For all of the functions in the study, the minimum line length (assuming a starting line length of 1) after correction is approximately 1; this allows us to simplify D_k as

$$D_k = \text{maximum line length in plot } k$$

Without any correction for the sine illusion, this factor is, like Tufte's lie-factor, equal to one. Values above one indicate that at least in some areas of the curve line segments are extended.

We compute this quantity for each sub-plot in each stimulus presented to the participant. The participant's choice therefore provides us with an estimate of what value of D constitutes the most consistent line length (out of the set shown). As each set of 6 plots is not guaranteed to contain a plot with $w = 0$, corresponding to constant length, choosing a plot with $D = 1.4$ indicates more distortion if there is a sub-plot with $w = 0$ ($D = 1$) present than if least distorted sub-plot present has $w = 0.4$ ($D = 1.2$ for plots of $y = \sin(x)$). Correcting for this bias, the set of $\{D_{\cdot\cdot k}\} = \{D_{\cdot\cdot 1}, ..., D_{\cdot\cdot 6}\}$ that is available to choose from produces an estimate of the overall psychological 'lie factor' as the ratio of the distortion of the chosen plot and the smallest distortion presented in set j:

$$P_{ij} = D_{ijk} / \min_{1 \leq k \leq 6} D_{ijk} \tag{5}$$

for each plot k and each participant i. This normalization does conservatively bias the results, effectively shrinking the effect size we observe, but without the normalization we would be biasing results in favor of a significant finding. Furthermore, while we could normalize relative to the maximal distortion, most stimuli sets contain large distortions values ($D > 1.85$; for $y = \sin(x)$), while more than half of the stimuli sets do not contain $D \approx 1$. This favoring of larger distortion values is due to choices of w around $w = 0.8$, which was suggested in the pilot data to be the commonly preferred weight value.

Whenever the set of available plots contains an uncorrected plot, the psycholgical lie factor P is equal to D; otherwise, i.e. no uncorrected plot is presented, the psycholgical lie factor P is less than the distortion D. This transformation is a conservative approach to estimating the lie factor, but allows us to show a variety of scaled transformations and estimate the effect with more granularity around $w = 0.8$.

By considering each participant's answers for the plot with the most consistent line length, we obtain an estimate of the psychological distortion from the sine illusion on an individual level. Estimating distortion factors for each participant facilitates comparison of these estimated values to determine whether the illusion is a product of an individual's perceptual experience or whether there possibly is an underlying perceptual heuristic for the sine illusion common across the majority of participants. If the illusion is a learned misperception rather than an underlying perceptual "bug", we would expect there to be considerable variability in the estimated individual lie factor P_i for each unique participant i, $1 \leq i \leq 123$, as it is likely that personal experience varies more widely than perceptual heuristics and their underlying neural architecture.

Each set of w values as defined in table I corresponds to a value of P as defined in equation 5. We test for a set of discrete values of w, which is reflected directly in the number of different values of P we can observe. This approach allows us to use a finite set of stimuli for testing, so that we can explicitly control the range of w displayed in each set of plots.

We model the theoretical psychological lie factor as θ, a continuous, unobservable population parameter, and individual psychological lie factors θ_i which represent the psychological distortion present due to the sine illusion in an individual. We can observe P_{ij} for each participant i and plot j. P_{ij} values are discrete (the set of possible P_{ij} values depends on the plot j), but provide information about the continuous θ_i for each participant (and the overall population value θ). We use a Bayesian model to account for the discrete nature of the observable values and the continuous unobservable parameter values.

Plots used in the experiment have factors P ranging between 1 and 2.5; we use a truncated normal data model for participant i viewing plot j, with $P = p_{ij} \sim N(\theta_i, \sigma)$ and independent flat prior distributions $\pi(\theta) = 0.4$ and $\pi(\sigma) = 2.5$ for $\sigma \in [.1, .5]$. These prior distributions $\pi(.)$ represent our expectations of the values of θ and σ before the experiment; assigning them constant values signals that we had little usable knowledge about the joint or marginal distributions of

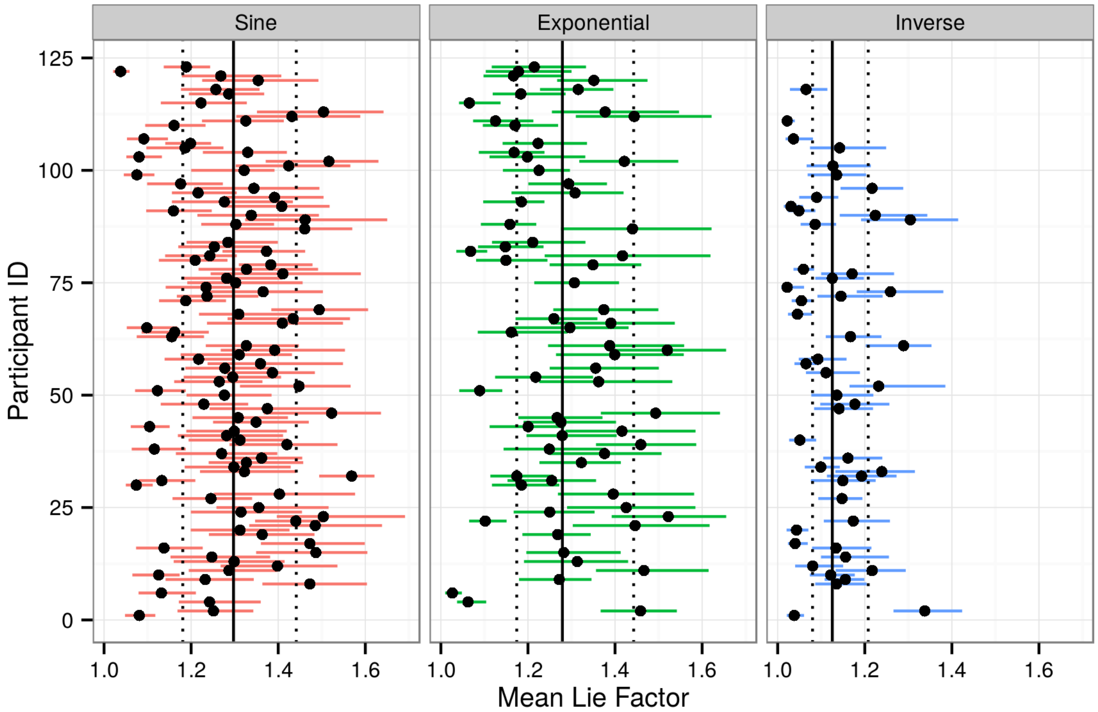


Fig. 12. 95% posterior predictive intervals for θ_i, calculated for each stimulus type. Vertical lines indicate the median estimate of the overall θ with a 95% credible interval.

θ and σ before the experiment was conducted. Using Bayesian estimation, we then obtain posterior distributions for θ_i and θ, the individual and overall mean lie factors. We are not particularly interested in the actual values of σ, but the additional parameter allows us to better estimate possible values for θ.

2.6 Results

the choice of names θ and θ_i suggests that θ is the population mean, while θ_i is the individual mean, and should therefore be centered around the population mean. everything but the last statement holds ... which makes figure **??** look so odd, because the individual densities are so different. Could we have one plot that shows the distribution of theta for all of the different functions, and then another plot like fig **??** for the individual densities separately? The individual densities suggest that there are at least two groups of people, and the overall population solution looks like an awkward fit between those two groups that doesn't capture either group adequately ... sorry ... I don't want to slam the statistics, it's just what it looks like to me. ... and I just noticed something else - are you sure that the lines show densities? The area under the curves don't seem to add to the same value. The population densities look like they encompass a larger area. Admittedly, the Bayesian approach is interested in $\pi(\theta)$, but I really think that the θ itself might be more informative - and the θ_i might then actually scatter around the population mean ... you can tell that I'm still struggling with this.

Figure 12 shows 95% posterior predictive credible intervals for θ_i estimated for each participant and mean function.

These intervals suggest that overall, the θ_i are similar across individuals. Very few (5 each for $y = \sin(x)$ and $y = \exp(x)$, 16 for $y = 1/x$) of the intervals overlap the region (1, 1.05), which corresponds to an "acceptable" lie

Table IV. Credible intervals for the overall θ for exponential, inverse, and sine stimuli.

Function	95% Credible Interval for θ	Median
Sin	(1.06, 1.23)	1.14
Exp	(1.2 , 1.46)	1.3
Inv	(1.17 , 1.4)	1.29

factor according to Tufte. This indicates significant distortion for most participants in our experiment, and the marked overlap of the intervals for each participant provides evidence consistent with a common magnitude of distortion. This suggests that there may be some common psychological strategy that is misapplied to the perception of these stimuli.

Comparison of the Preferred Stimuli. Estimates of $\hat{\theta} = E[\theta]$ for each function are 1.31, 1.29, and 1.14 respectively for exponential, inverse, and sine functions, suggesting a similar psychological distortion even for very different functions, though it seems as if the inverse function causes somewhat less distortion, possibly because the correction factor is not as proportionately large. Credible intervals can be found in Table IV. As all three of the credible intervals exclude 1.05, there is evidence that a psychological distortion is occurring; that is, there is evidence of a significant psychological lie factor. In addition, the method of adjusting the estimated lie factor we have used here is conservative; it is likely that because most stimuli contain a sub-plot with $w \geq 1$ our estimate of θ for the inverse function is low (as the lie factor for fully corrected inverse plots is greater than the corresponding lie factors for the other two functions used in this experiment).

The estimated weight values corresponding to these θ are shown in figure 13. In all three cases, the experimentally-corrected plots appear less distorted than the uncorrected plots.

This experiment has demonstrated that the sine illusion results in misperception of graphically presented data. In particular, participants tend to see uneven line length when lines are even while missing uneven line length due to the illusion's effect.

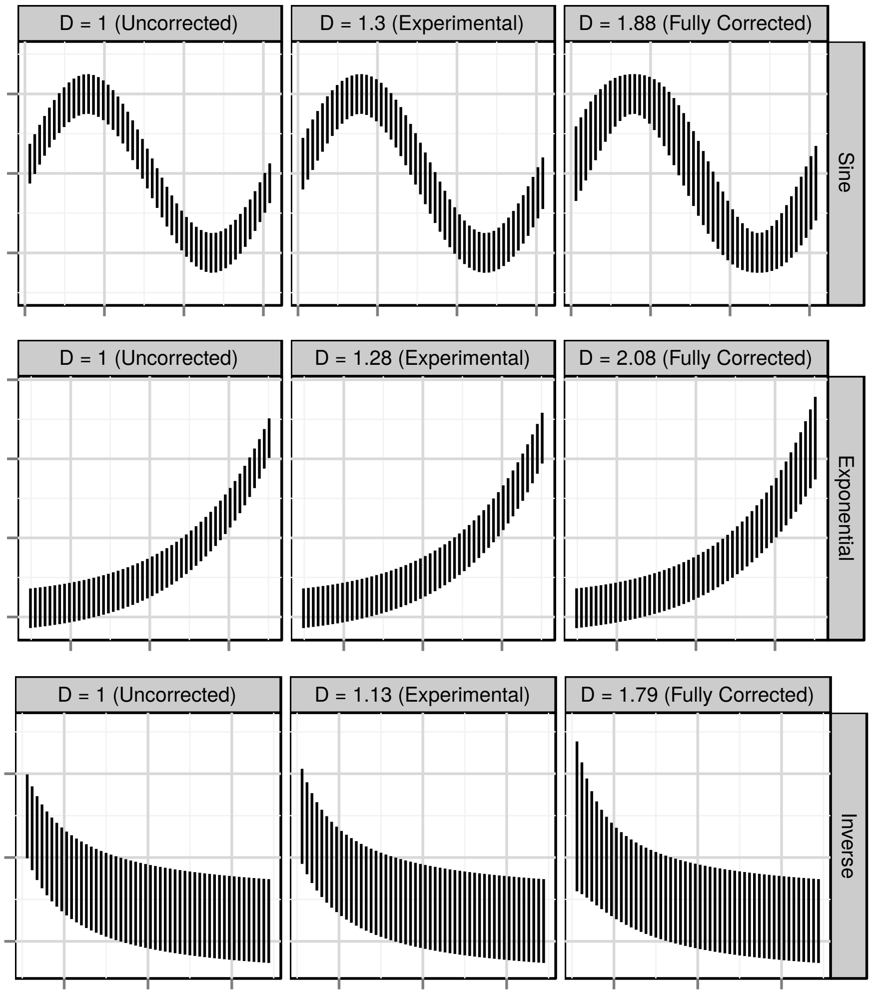


Fig. 13. Uncorrected, experimentally corrected (according to the median value of θ), and fully corrected stimuli for all three underlying functions used in the experiment. The corrected value shown here is equivalent to the distortion factor D defined as $D = \ell_{max}/\ell_{min}$.

3. CONCLUSIONS

The sine illusion arises from misapplication of three-dimensional context to a two-dimensional stimulus which results in nearly unavoidable perceptual distortions. These distortions impact the inferences made from charts which are similar to three-dimensional figures, even when we are not consciously aware that this context exists; the only immunity to the illusion we have found is for those who have never developed binocular depth perception. We have estimated that the illusion produces a distortion of about 135%. This distortion occurs entirely between the retinal image and the mental representation of the object; it is not due to the chart, rather, it is an artifact of our perceptual system.

As Tufte advocated for charts that showed the data without distortion, our goal is to raise awareness of perceptual distortions that occur within the brain itself due to misapplied heuristics. While applying corrections to the data to remove these distortions is somewhat radical, the persistence of the illusion despite awareness of its presence presents a challenge to those seeking to display data visually. In addition, many graph types can induce this illusion (scatterplots, ribbon plots, parallel sets plots), so avoiding a specific type of graph is not an effective solution. The best solution to this problem is to raise awareness: to demonstrate that optical illusions occur within information graphics, and to understand how these illusions arise.

REFERENCES

AHLUWALIA, A. 1978. An intra-cultural investigation of susceptibility to "perspective" and "non-perspective" spatial illusions. British Journal of Psychology 69, 233–241.

AMER, T. 2005. Bias due to visual illusion in the graphical presentation of accounting information. Journal of Information Systems 19, 1, 1–18.

CLEVELAND, W. S. AND MCGILL, R. 1984. Graphical perception: Theory, experimentation, and application to the development of graphical methods. Journal of the American Statistical Association 79, 387, pp. 531–554.

CLEVELAND, W. S. AND MCGILL, R. 1985. Graphical perception and graphical methods for analyzing scientific data. Science 229, 4716, 828–833.

DAY, R. H. AND STECHER, E. J. 1991. Sine of an illusion. Perception 20, 49–55.

GREGORY, R. 1968. Perceptual illusions and brain models. Proc. Roy. Soc. B 171, 279–296.

HENSON, D. B. AND WILLIAMS, D. E. 1980. Depth perception in strabismus. British Journal of Ophthalmology 64, 5, 349–353.

HOFMANN, H. AND VENDETTUOLI, M. 2013. Common angle plots as perception-true visualizations of categorical associations. Visualization and Computer Graphics, IEEE Transactions on 19, 12, 2297–2305.

HOLOPIGIAN, K., BLAKE, R., AND GREENWALD, M. J. 1986. Selective losses in binocular vision in anisometropic amblyopes. Vision research 26, 4, 621–630.

HUBEL, D. H. AND WIESEL, T. N. 1970. The period of susceptibility to the physiological effects of unilateral eye closure in kittens. The Journal of physiology 206, 2, 419–436.

KOSARA, R. AND ZIEMKIEWICZ, C. 2010. Do mechanical turks dream of square pie charts? In Proceedings BEyond time and errors: novel evaLuation methods for Information Visualization (BELIV). ACM Press, 373–382.

LARKIN, J. H. AND SIMON, H. A. 1987. Why a diagram is (sometimes) worth ten thousand words. Cognitive science 11, 1, 65–100.

MAYER, R. E. AND SIMS, V. K. 1994. For whom is a picture worth a thousand words? extensions of a dual-coding theory of multimedia learning. Journal of educational psychology 86, 3, 389.

PARKER, A. J. 2007. Binocular depth perception and the cerebral cortex. Nature Reviews Neuroscience 8, 5, 379–391.

PENROSE, L. S. AND PENROSE, R. 1958. Impossible objects: A special type of visual illusion. British Journal of Psychology 49, 1, 31–33.

PLAYFAIR, W. 1786. Commercial and Political Atlas. London.

POULTON, E. 1985. Geometric illusions in reading graphs. Perception & psychophysics 37, 6, 543–548.

ROBBINS, N. 2005. Creating More Effective Graphs. Wiley.

SCHONLAU, M. 2003. Visualizing categorical data arising in the health sciences using hammock plots. In Proceedings of the Section on Statistical Graphics (JSM '03). American Statistical Association.

SECKEL, A. 2007. Masters of Deception: Escher, Dal & the Artists of Optical Illusion. Sterling.

TUFTE, E. 1991. The Visual Display of Quantitative Information 2 Ed. Graphics Press, USA.

VANDERPLAS, S. AND HOFMANN, H. 2014. Signs of the sine illusion - why we need to care. Journal of Computational and Graphical Statistics.

ZHANG, J. AND NORMAN, D. A. 1994. Representations in distributed cognitive tasks. Cognitive science 18, 1, 87–122.