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

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

1. INTRODUCTION

One of the basic principles of visual graphics is that the graphic should accurately reflect the data. Tufte's "Lie Factor", for instance, is calculated based on the ratio of the effect size shown in the graphic to the effect size in the data [Tufte 1991]. If there is a systematic difference, the lie factor will be notably different from 1 (values between .95 and 1.05 are typically acceptable). While Tufte's lie factor is an effective measurement of the transition from data to graphics, it is not as effective at measuring the transition from graph to the brain. Ideally, graphics would not only represent the data accurately, but would also allow readers to draw accurate conclusions from the graph. While in many cases these two goals can be achieved simultaneously, there are instances in which human perception produces a psychological distortion that can interfere with the interpretation of graphics, even if the graphics are accurate portrayals of the numerical data. In this paper, we examine situations in which low-level human perceptual processes interfere with making accurate judgements from displays and suggest an experimental methodology for estimating the psychological "lie factor" of a graphic.

2. 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 reasonably accurate 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 Necker Cube shown in figure 1. Additionally, the same three-dimensional object often has infinitely many two-dimensional representations, for instance, when viewed from different angles. Optical illusions that demonstrate this phenomenon are the so-called "impossible objects" such as the Penrose triangle. In the case of the Necker cube, the two different figures are equally salient, and thus the brain does not prefer one interpretation over the other and instead continuously switches between interpretations. In many other instances
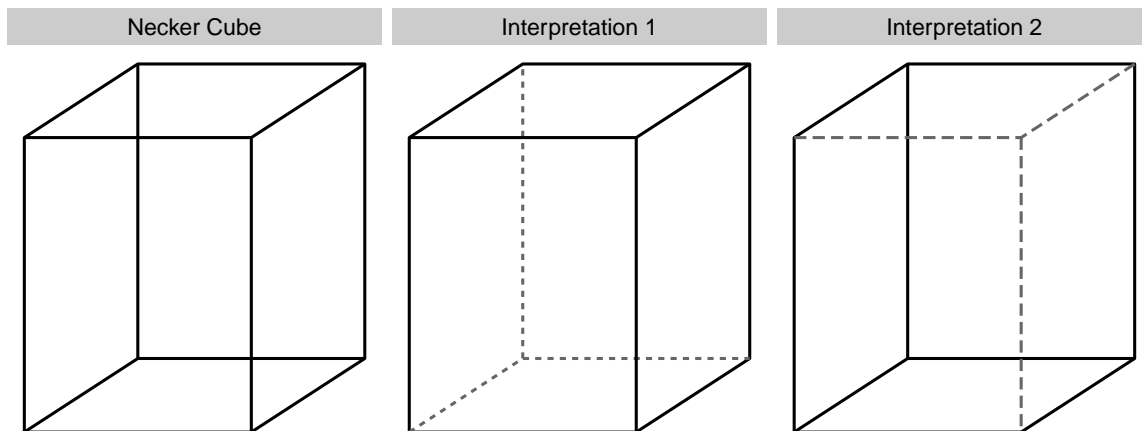

Fig. 1.   The Necker Cube is a so-called "ambiguous object" because two different transparent objects could produce the same retinal image (and thus the same perceptual experience). Commonly, the image seems to transition instantaneously from one possible mental representation to the other.[Gregory 1968]
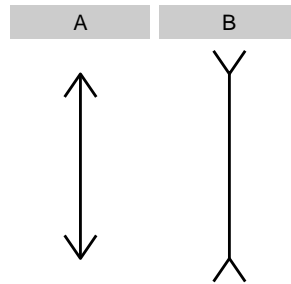
Fig. 2.   The Müller-Lyer illusion. The central segment of figure A is perceived as shorter than the central segment of figure B, even though the two are actually the same length.

experience with the real world informs the choice between multiple possible three-dimensional objects which form the same two-dimensional picture. 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 the picture 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.

One explanation for the Müller-Lyer illusion is that the brain interprets the ambiguous lines as a common three-dimensional object: 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 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, even though the retinal size is identical for the two objects. This inferred "depth cue" [Gregory 1968] is reasonably consistent across individuals. 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. The advantages of this approach in the real world are considerable [Gregory 1968], as pictures of real objects are seldom ambiguous and this strategy allow for high performance with limited neural bandwidth. Figure 3 shows not only the real-world context, but that the perspective cues which contribute to the illusion allow for an accurate neural representation of the object in context.

Fig. 3.   Real-world context that gives rise to the Müller-Lyer illusion

3.   THREE DIMENSIONAL CONTEXT OF THE SINE ILLUSION

While the classic Müller-Lyer illusion is seldom a factor in statistical graphics, 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) is one example of this phenomenon which occurs frequently in statistical graphics. Like the Müller-Lyer illusion, it is pervasive and very difficult to "un-see" or mentally correct. Figure 4 shows the sine illusion in its original form, and figure 5 shows the same illusion in one of the charts from Playfair's Statistical Atlas [Playfair 1786]. This illusion, which is also known as the line-width illusion, has also been documented in parallel sets plots [Hofmann and Vendettuoli 2013].

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 readily confirm 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. 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-dimensional image that may influence the perceptual context. One of these contexts is shown in figure 6, generated from the same function shown in the two-dimensional

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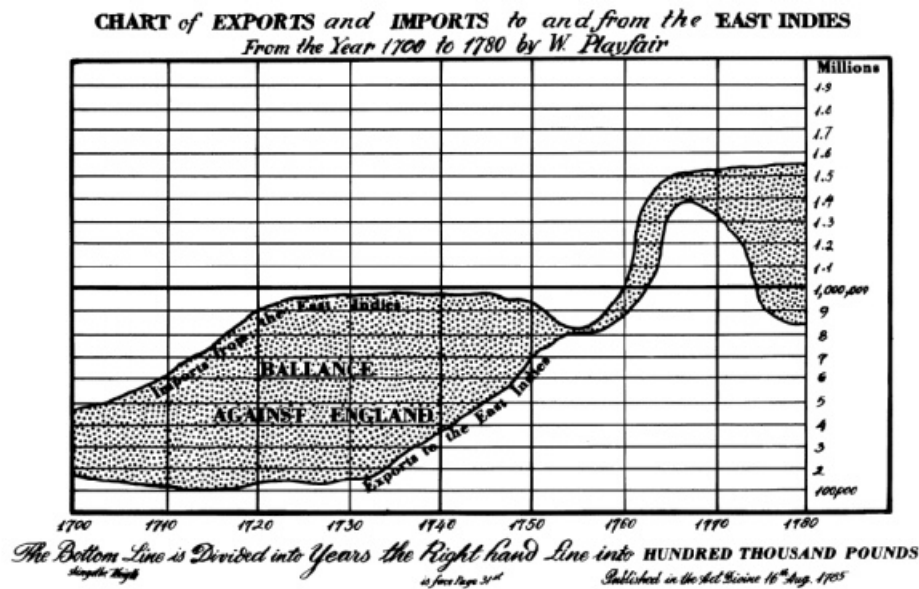
Fig. 5. The sine illusion in Playfair's graph of trade between the East Indies and England, 1700-1780. The trade balance in 1763 does not appear to be the same size as that in 1745, though the vertical distance is approximately the same.

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. One main difference between the first three dimensional image shown in figure 6 and the original image is that the lines connecting the top

Finite Vanishing Point   Near–infinite Vanishing Point

Fig. 6.   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.

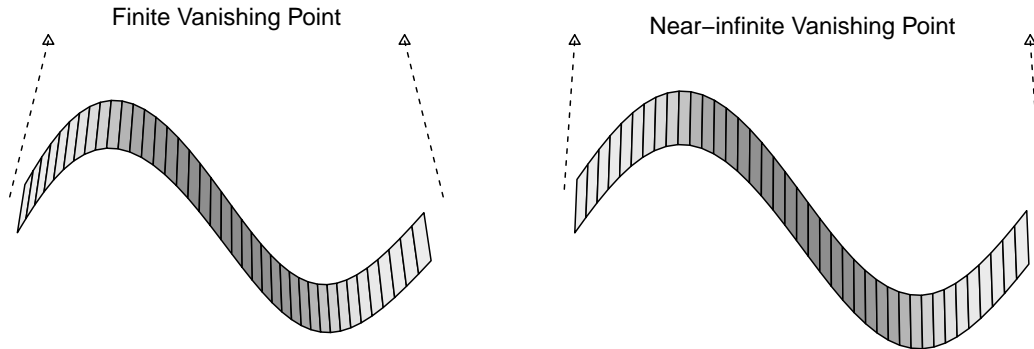and bottom sections of the curve are slightly angled in the three-dimensional version; this is due to the perspective projection used to create the image and the corresponding angles of rotation chosen such that the entire surface is visible.

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 on figure 4. The second image in figure 6 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 Muller-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.

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.

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

4. MEASURING THE PSYCHOLOGICAL LIE FACTOR EXPERIMENTALLY

4.1 Study Setup

In order to determine the amount of the psychological distortion, participants were presented with different stimuli consisting of six subplots, similar to figure 7. These sets of charts were constructed such that line lengths along a curve were varied to different degrees to counteract the illusion. Figure 8 shows the amount of line correction used in each of the sub-plots above. 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.

The amount of correction in each sub-plot was chosen such that each stimuli set contained a selection of curves corrected to various degrees using correction factors determined previously in an internal pilot study. The correction factor we use for extending the length of the line segment at loctaion x is given as

$$(1-w) + w \cdot \left(1/\cos\left(\tan^{-1}(|f'(x)|)\right)\right)$$
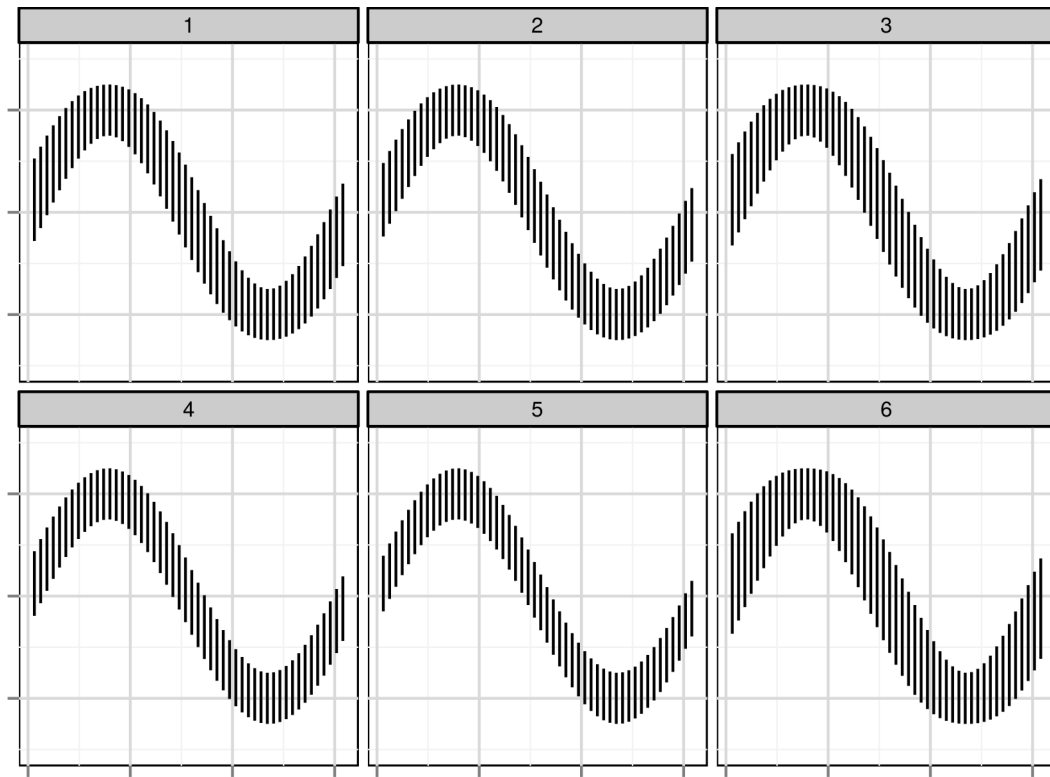
Fig. 7.   One of the graphs presented to participants through Amazon Mechanical Turk. Figure 8 shows the actual differences in line lengths.
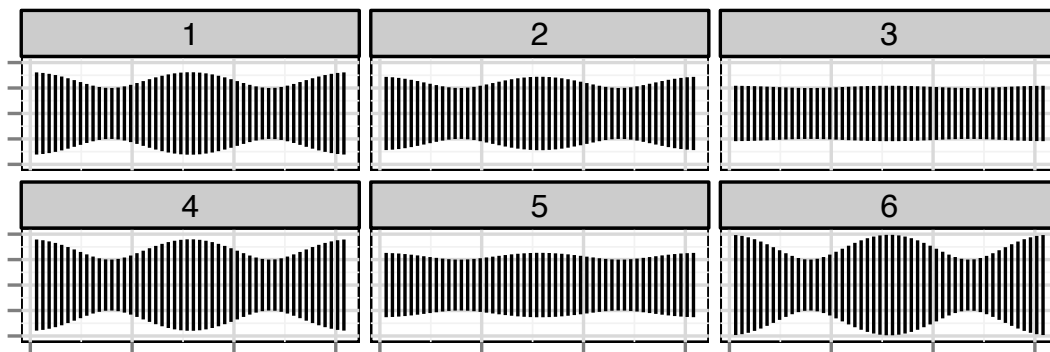
Fig. 8.   Line lengths (without the trend) for figure 7, demonstrating the effective lie factor for each sub-plot.

where w is varied between 0 (for no correction) and 1.5 (strong over-correction). From a pilot study, we suspected values around $w = 0.8$ to be promising, but did not know whether this would generalize to those outside of the statistical graphics community. This led to twelve stimuli sets with correction factors as shown in table I. The difficulty level of each of the twelve stimuli sets was determined by the similarity of the six plots shown.

Three underlying mean functions were tested using these weight values: $\sin(x)$, $\exp(x)$, and $1/x$.

Table I. Ordered weight factors and corresponding line factors for the sine curve stimuli sets.

set	difficulty	Weight (w)						Lie Factor (D)					
		sub-plot						sub-plot					
		1	2	3	4	5	6	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

Each participant was presented with eleven sets of graphs, consisting of one "easy" test chart, 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. Plots were shown to participants in random order. The test chart consists of a set of six sine curves with a very low level diffcilty level, and is used as an introduction to the testing procedure.

Though participants were asked to participate in 11 total trials, some participants continued to provide feedback beyond the eleven trials required to receive payment through Amazon. For any subsequent responses we randomly selected one of the 32 possible stimuli. This approach allowed us to collect some data in which a single participant provided responses to all three underlying functions.

4.2 Demographics

Participants for the study were recruited through the Amazon Turk web service. Responses from 128 users at 123 unique IP addresses were collected.

In order to more accurately identify unique individuals, data were grouped by IP address rather than by user-entered identification code, as it is common for individuals to have multiple accounts and somewhat less likely for two Amazon Mechanical Turk participants to share the same IP address and complete the same study in a reasonably short window. We treated responses from the same IP addresses as additional responses from the same user.

4.3 Analysis

Psychological "Lie Factor". As the strength of the correction varies across the horizontal range of the curve, we quantify the psychological distortion as the ratio of the maximum line length to the minimum line length for each plot j: $D_j = l_{max}/l_{min}$, and denote the chosen ratio as D^*. 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 $w = 0.4$ ($D = 1.2$). Correcting for this bias, the set of $D^* = \{D_1, ..., D_6\}$ that is available to choose from produces an estimate of the overall psychological 'lie factor' as

$$P = D^* / \min_{1 \leq j \leq 6} D_j \tag{1}$$

for each plot and each participant. That is, P is the ratio of the lie factor of the chosen plot to the smallest lie factor available in the set of available plots. When the set of available plots contains an uncorrected plot, $P = D$; when an uncorrected plot is not presented, $P < D$, since each $D_j >= 1$, $j = 1, ..., 6$. 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 can 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 is a possible 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 1. We test for only 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. To mathematically model a continuous quantity (the real domain of possible P values) using discrete data, we employ a Bayesian approach to model an overall psychological lie factor θ and individual participant lie factors θ_i.

Plots used in the experiment have factors P ranging between 1 and 2.5, so we can use a truncated normal data model for participant i viewing plot j, with $P = p_{ij} \sim N(\theta_i, \sigma)$ and independent flat priors $\pi(\theta) = 0.4$ 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 indicates that we had little useable knowledge about the joint or marginal distributions of θ and σ before the experiment was conducted. Using Bayesian estimation, we can 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 is a useful tool to better estimate possible values for θ.

4.4 Results

The posterior density of θ for each function is shown in figure 9, along with separate posterior densities for each individual θ_i. θ is reasonably similar for all three functions, suggesting that while function type may moderate the size of the effect, the illusion occurs regardless of function type. Individual curves have different variability due to the number of trials completed, and are necessarily more spread out as there is less data with which to estimate the individual posterior distributions.

On an individual level, figure 10 shows the posterior density for θ_i for four of the participants who completed at least 6 trials in each category. While in many cases, the most probable θ_i is similar across trials, individuals do seem to have been somewhat more affected by the illusion when the underlying function was sinusoidal, though this may reflect a discrepancy in the number of trials rather than a stronger illusion. Alternately, as the illusion depends on variable slope, it is possible that the monotonic exponential and inverse stimuli induced a weaker three-dimensional context.

In order to appropriately compare intervals for each participant's θ_i, we simulated 11 new "data points" from our model (thus enforcing a uniform 11 trials per participant for each function type) to get a single new estimate of $\hat{\theta}_i$. For each participant, we generated 100 of these $\hat{\theta}_i$ and used these simulated values to calculate the 95% credible intervals shown in figure 11. These intervals will allow us to consider the variability in θ_i due to participant preference rather than the number of trials a participant completed during the study. Removing this additional variability provides us with the opportunity to consider whether the sine illusion stems from an individual's perceptual experiences or from a lower-level perceptual heuristic.

Posterior predictive intervals for θ_i as shown in figure 11 suggest that overall, the θ_i are similar across individuals. Very few of the intervals overlap the region (1, 1.05), which corresponds to an "acceptable" lie factor according to
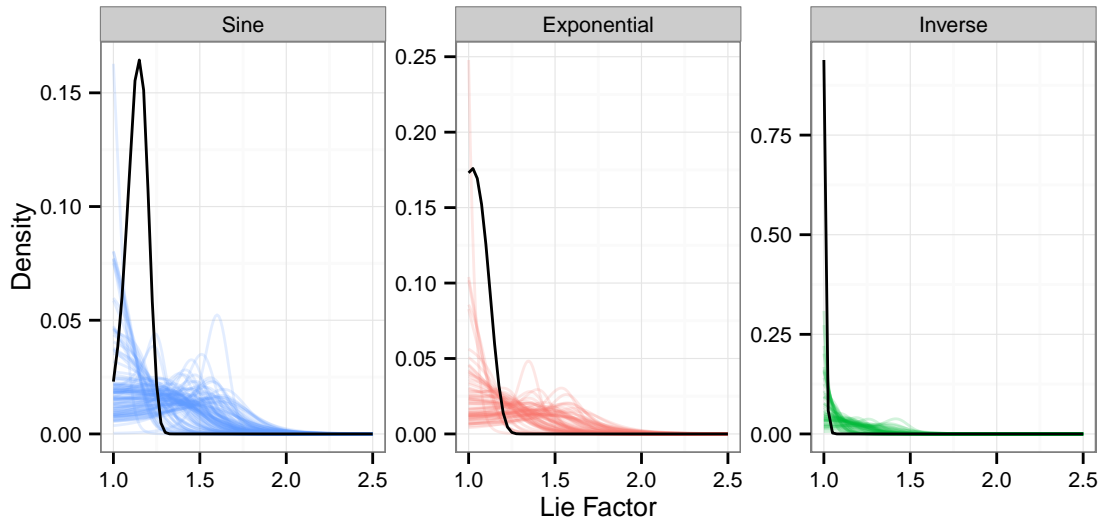
Fig. 9. Estimated densities for θ_i, shown in color, with the estimated overall density for θ shown in black. Individuals have extremely similar posterior distribution of θ_i, and even different functions have similar $\hat{\theta}$, suggesting a common underlying mental distortion.
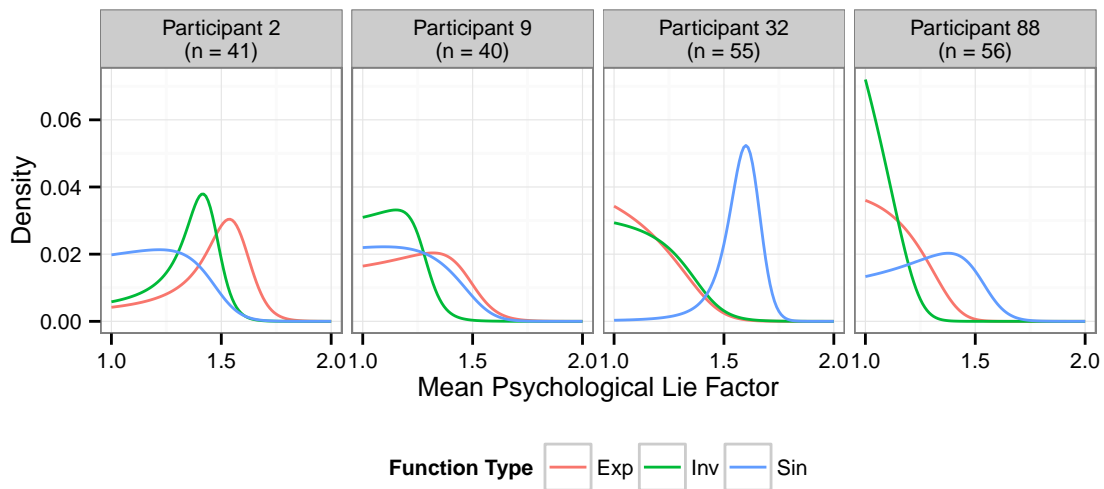

Fig. 10. Posterior distributions for θ_i for four of the participants who completed at least 6 trials of each of the three function types.

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.
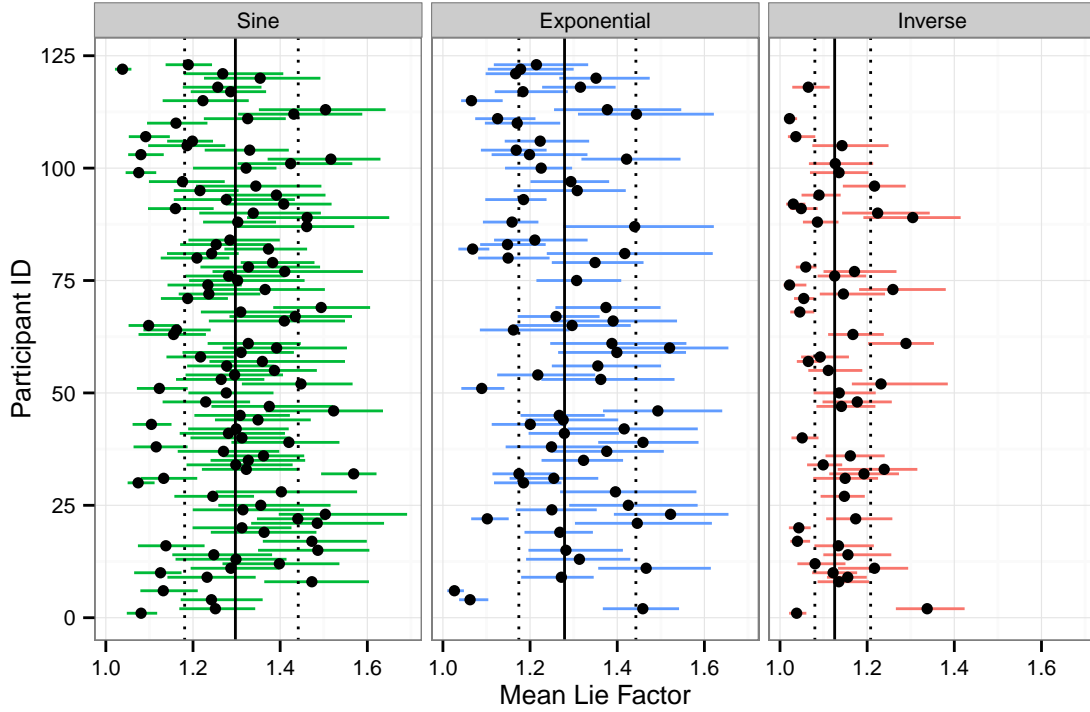
Fig. 11. 95% posterior predictive intervals for θ_i, calculated for each stimulus type. Vertical lines indicate the median estimate of the overall θ with a 95% credible interval.

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

Function	95% Credible Interval for θ	Median
Sin	(1.1808, 1.4415)	1.2973
Exp	(1.1742, 1.4428)	1.2791
Inv	(1.08, 1.208)	1.1255

Comparison of the Preferred Stimuli. Estimates of $\hat{\theta} = E[\theta]$ for each function are 1.29, 1.13, and 1.3 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 II. 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 optimal weight values corresponding to these θ are shown in figure 12. In all three cases, the optimally corrected plots appear less distorted than the uncorrected plots.
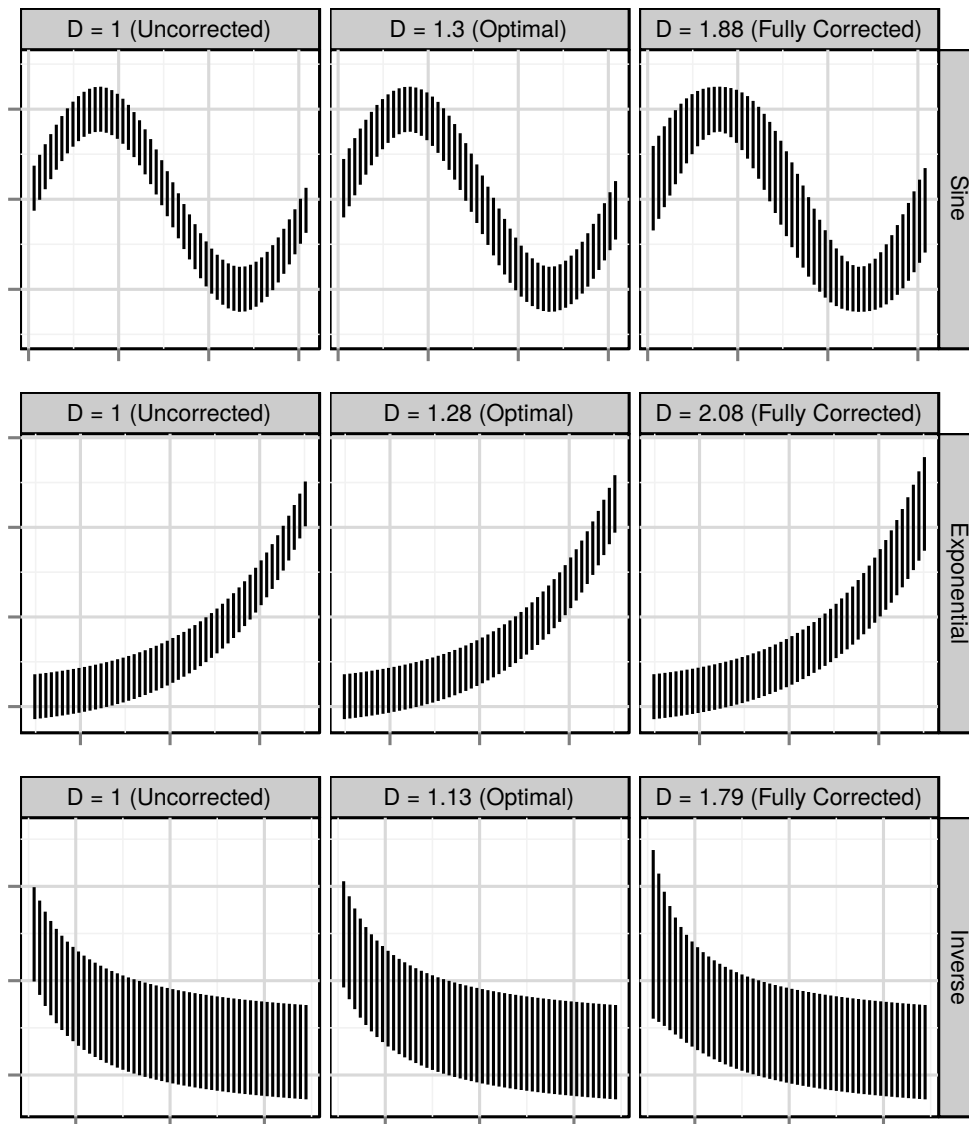
Fig. 12. Uncorrected, optimally corrected (according to the median value of θ), and fully corrected stimuli for all three underlying functions used in the experiment. The corrected value shown here is equivalent to the distortion factor D defined as $D = \ell_{max}/\ell_{min}$.

5. CONCLUSIONS

The sine illusion arises from misapplication of three-dimensional context to a two-dimensional stimulus which results in nearly unavoidable perceptual distortions that impact the inferences made from graphics. We have estimated that the illuson 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 graphics 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 statistical 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.

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.

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.

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

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