Common Angle Plots as perception-true visualizations of categorical associations

Heike Hofmann, *Member, IEEE,* and Marie Vendettuoli

Abstract—Visualizations are great tools of communications - they summarize findings and quickly convey main messages to our audience. As designers of charts we have to make sure that information is shown with a minimum of distortion. We have to also consider illusions and other perceptual limitations of our audience. In this paper we discuss the effect and strength of the line width illusion, a Müller-Lyer type illusion, on designs related to displaying associations between categorical variables. Parallel sets and hammock plots are both affected by line width illusions. We introduce the common-angle plot as an alternative method for displaying categorical data in a manner that minimizes the effect from perceptual illusions. Finally, we present results from user studies as evidence that common angle charts resolve problems with the line width illusion.

Index Terms—Linewidth illusion, Data Visualization, High-dimensional Displays, Parallel Sets, Hammock Plots, Müller-Lyer Illusion.

I. INTRODUCTION

GRAPHICS are tools of communication: the purpose of a good chart is to summarize quantitative information and present it in a way that enables readers to extract the main message easily and with a minimum of distortion. Tufte coined the amount of distortion the *Lie-Factor* (1, p. 57–69) and defined it as the ratio of the effect size displayed and the size of the same effect in the data. A ratio of one indicates no distortion, values less than one indicate a visual underrepresentation of an effect, while values above one indicate a corresponding overrepresentation. Three-dimensional depth effects, in particular, are prone to large Lie-Factors.

Perceptual limitations of our audience are another source of distortions. Illusions give rise to a distortion similar to a Lie-Factor effect. However, these distortions are not directly measurable in the actual representation: a physical quantity drawn in the chart might perfectly reflect the size of an effect in the data, but if an observer intuitively interprets another property of the geometrical object, this leads to distorted, if not altogether wrong, conclusions. As designers of charts we have to be aware of the innate or culturally imposed perceptual limitations of our audience. Good graphical designs minimize and counteract these limitations. Since we cannot measure effects from illusions directly, we have to employ usability studies to provide

H. Hofmann is faculty member in the Department of Statistics, Iowa State University, Ames, IA, 50014 USA e-mail: hofmann@iastate.edu.

M. Vnedettuoli Graduate Student in the HCI and BCB program of Iowa State University and member of the working group at the USDA in Ames, IA 50014.
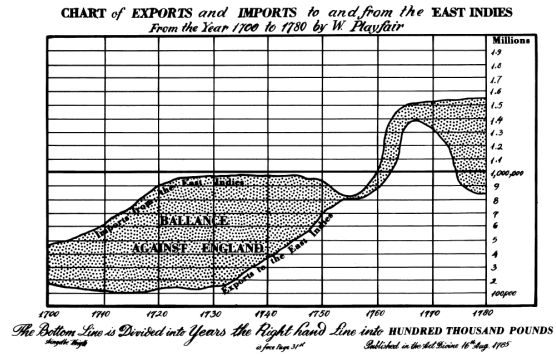
Fig. 1. Playfair's chart from the Commercial and Political Atlas (1786) showing the balance of trade between England and the East Indies. In which years was the difference between imports and exports the highest?

empirical evidence in support of underlying metaphorical models or for identification of unexpected or unpredicted effects from perceptual illusions.

In this paper we discuss the effect and strength of the *line width illusion* on designs related to displaying associations between categorical variables. We will focus specifically on the use of parallel coordinates (2; 3; 4) for categorical data and the perceptual implications this has.

Parallel coordinates were introduced to visualize large numbers of numeric variables. Recent adaptations of parallel coordinates (5; 6; 7; 8; 9) and other, frequency-based, methods to display categorical data, such as mosaic plots (10; 11; 12; 13), have addressed perceptual constraints (14) at a theoretical level, but lack reporting of human subjects testing as validation.

We will first show how the line width illusion affects parallel coordinates for categorical data, and in a second step introduce the *common-angle plot* as an alternative method for displaying categorical data in a manner that minimizes the effect from perceptual illusions. The display preserves properties of parallel coordinates, such as the potential to visualize a large number of dimensions simultaneously, but also presents frequency information.

Finally, we present results from user studies as evidence that common angle plots resolve the problem of the line width illusion.

II. ILLUSIONS AND DISPLAYS

A. Line width illusion

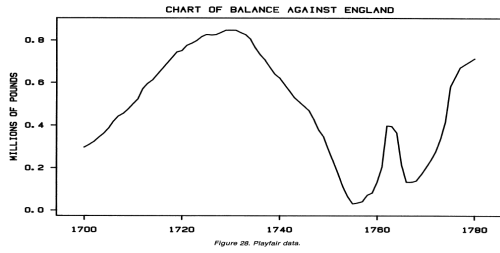An example of the *line width illusion* is displayed in figure 1.

Fig. 2. Difference between exports and imports from England to and from the East Indies in the 18th century – the steep rise in the difference around 1760 comes as a surprise to many viewers of the raw data in figure 1.

This chart displays the balance of trade between England and the East Indies as shown by William Playfair in his Commercial and Political Atlas, 1786 (15; 16). One purpose of this chart is to demonstrate the difference between imports and exports in a particular year and its pattern over that time frame. The difference in exports and imports is encoded as the vertical difference between the lines. When observers are asked to sketch out the difference between exports and imports (14), they very often miss the steep rise in the difference between the lines in the years between about 1755 and 1765. Figure 2 shows the actual difference between imports and exports.

This phenomenon is known and widely discussed in statistical graphics literature (14; 1; 17; 18). It is due to our tendency to assess distance between curves as the minimal (orthogonal) distance rather than the vertical distance – see sketch 4 for a visual representation of both.

In the perception literature, this phenomenon is known as part of a group of geometrical optical misperceptions of a context-sensitive nature classified as Müller-Lyer illusions (19). Interestingly, there seems to be a general agreement that this illusion exists, but a quantification of it is curiously absent from the literature.

While we see the type of chart as shown in figure 1 proposed by Playfair quite commonly, particular in election years – where these kind of charts are used to enable comparisons of support for several candidates, the recommendation from the literature is to avoid charts in which the audience is asked to do visual subtractions, and show these differences directly.

However, the line width illusion is not restricted to this situation only. We next discuss how other charts, such as the parallel sets plots (5), are affected by it.

B. Parallel sets

Parallel sets (parsets) have been introduced by R Kosara (5) as one way to visualize multivariate categorical data. Since their initial publication, they have spread to mass media outlets: see e.g. "How to understand risk in 13 clicks," BBC (20).

While retaining the ability to visualize a large number of dimensions simultaneously that is the parallel coordinates' hallmark trade, parallel sets introduce the frequency scale that is a well-known feature of other categorical displays

such as barcharts or mosaic plots (10; 11; 12; 13). Initially, frequencies of categories were displayed as stacked boxes; in later versions of parallel sets the boxes are reduced to simple lines (21). Various implementations of parallel sets exist besides the original Java version of Eagereyes (22), e.g. J Davies introduced a d3 (23) version in (24). The implementation of parsets-style displays in this paper is done in the software R (25). It retains the original box placeholders and does not enforce a tree hierarchy via color grouping. Section V gives details on the implementation, section III-C2 discusses dealings with hierarchical structures in the data.
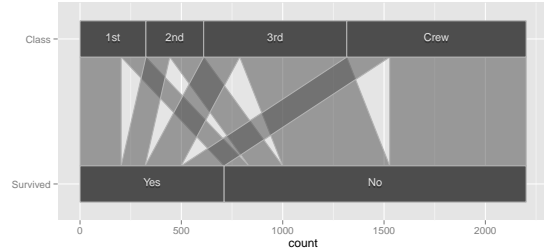

Fig. 3. Parallel sets plot showing the relationship between survival of the sinking of the HMS Titanic and class membership.

Figure 3 gives an example for a two dimensional parallel sets plot investigating the relationship between class status and survival on board the HMS Titanic (26). Class status is recorded as either crew member or passengers in first, second, or third class.

The top bar in figure 3 shows the variable Class. The bottom bar shows survival as yes and no. Between the bars lines are drawn to visualize the relationship between class membership and survival. Based on the number of survivor and non survivors these bands are drawn from each class, and their (horizontal) width is proportional to the number of people they represent.

This definition implies that lines are drawn under different angles, which makes the plots susceptible to the line width illusion. As an example, try to order the levels of variable 'class' according to the number of survivors they have. In order to do that, we need to evaluate the width of the bands between upper and lower bars in figure 3. The perceived width of the lines between the variables in figure 3 might (and, as the responses from the survey in section IV-B show, does) lead us to assume that 1st class passengers had the highest number of survivors, followed by 3rd class passengers and about even numbers of survivors in the 2nd class and the crew. However, the actual numbers are:

	Crew	1st	2nd	3rd
Survivors	212	203	118	178
Non-Survivors	673	122	167	528

This identifies an order of crew, first class, third class, followed by second class as correct. We will discuss next how this finding relates to the line width illusion and how this helps us in quantifying the strength of the illusion.

C. Strength of line width illusion

When visually evaluating lines of thickness greater than one, the line width illusion applies, only now the *edges* of a single line take on the role of the separate curves. As above, there is a strong preference of evaluating the width of lines orthogonal to their slopes as opposed to horizontally (see figure 4) needed for a correct evaluation of parsets-style displays.

Orthogonal w_o and horizontal w_h line widths are related – the orthogonal line width depends on the angle (or, equivalently, the slope) of the line:

$$w_o = w_h \sin \theta, \tag{1}$$

where θ is the angle of the line with respect to the horizontal line.
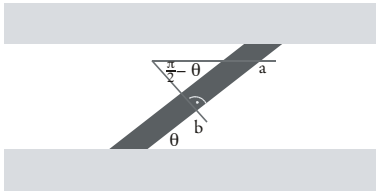

Fig. 4. Sketch of line width assessments: (a) is showing horizontal width, (b) shows width orthogonal to the slope. From the survey results in section IV-B we can see that observers associate line width more with orthogonal width (b) than horizontal width (a).

The perceived slope of a line very much depends on the aspect ratio of the corresponding plot – changing the height to width ratio of a display will change our perception of the corresponding line widths, if they are not adjusted for the slope (14). This finding is not new, but its strength on our perception is surprising, as can be seen in the example of figure 5. Again, survival and class membership on the Titanic is shown; the same parallel sets plot is shown twice in this figure, but with very different aspect ratios: in the plot on the left the number of surviving 3rd class passengers seems to be about twice as big as the number of survivors among crew members, whereas in the plot on the right the lines have about equal (orthogonal) width. Obviously, this is not due to a change in numbers.

For parsets-style displays, the audience has *area of the line segment* an alternate visual cue when evaluating frequencies. Because height (or width for a rotated display) of line segments is constant across the display, the width of a particular segment is proportional to its area. We can therefore employ area comparisons as a proxy or to augment line width evaluations. However, the literature suggests that this method of comparison is particularly prone to errors in two scenarios commonly seen in parallel sets: (1) extreme aspect ratios of the rectangular shape (27) and (2) when comparing rectangles rotated relative to each other (28). This incorrect perception and comparison of areas provides additional contextual evidence that reinforce and strengthen distortion introduced by the line width illusion.
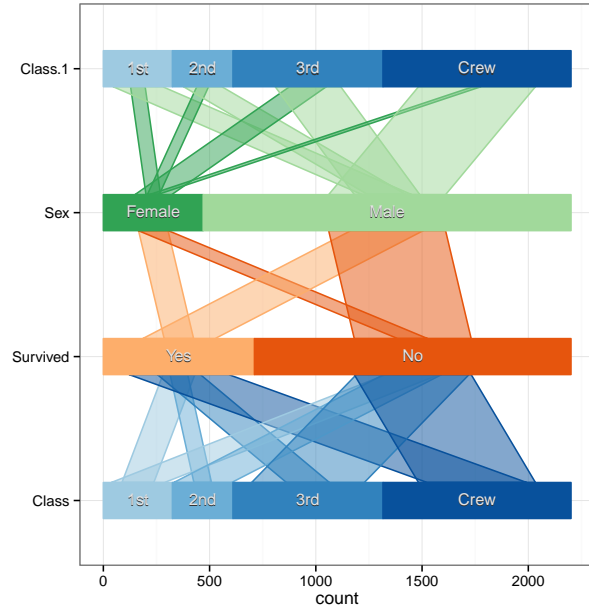

Fig. 6. Hammock plot of the relationship between Class and Survival on the Titanic.

Designers of good graphics have to accommodate changes in intuitive evaluation of the situation. While re-ordering categories and controlling the aspect ratio reduces the potential that audiences draw wrong conclusions from a plot, none of these measures is enough to avoid all perceptual pitfalls due to the line width illusion.

D. Hammock plots

Hammock plots, introduced by M Schonlau in (9), provide an alternative to parallel sets that is adjusted for the line width illusion. This is done by adjusting the – horizontal– line width by a factor of $\sin \theta$, as discussed in equation (1). This adjustment makes the perceived – orthogonal– line width to be proportional to the number of observations it represents. Figure 6 shows an example of a four dimensional hammock plot of the Titanic data. From top to bottom Class, Gender, Survival, and again Class are shown.

Similarly to the parallel sets plot, the bars are divided according to class membership numbers. Lines connect categories between neighboring variables, orthogonal line widths are representing the number of individuals in each combination. Unlike the parallel sets, the lines start from the middle of the bin and connect to the middle of the other variable's bins.

Here, we see that barely any women were in the crew, while male crew members make up the second largest contingent overall. Overall a few more men survived than women. Proportionally the situation is much different – a much higher percentage of women survived than men. We also see, that more first class passengers survived than not, while survival chances of second class passengers were about fifty-fifty. For third class passengers and crew members fewer members did survive than not.
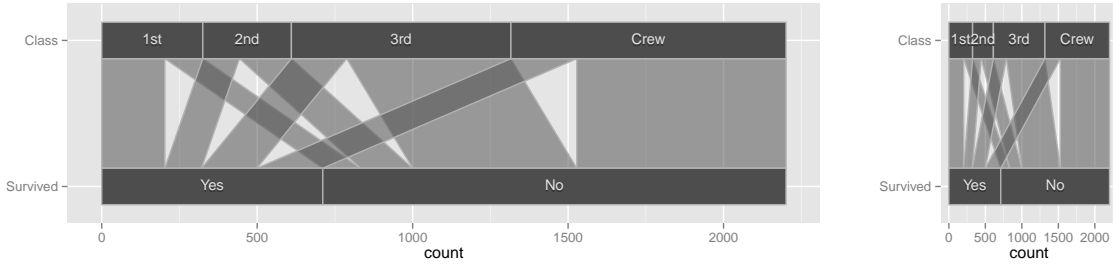
Fig. 5. Parallel sets plots of survival on the Titanic by class. Different aspect ratios seemingly change the thickness of line segments, compare e.g. number of survivors in 3rd class and in the crew.

As the adjustment of line widths is made with respect to the angle θ, which itself depends on the aspect ratio of a plot, we need complete control over these properties of the plotting device when constructing hammock plots – in our implementation (see below for details) we have dealt with this issue by fixing the aspect ratio. This is problematic in some situations, where the rendering has to be done without knowledge of the plotting device.

Another problem that arises in evaluating hammock plots is that if an observer focuses on horizontal line width the plots suffer from a *reverse line width illusion*: judging the number of survivors by class in figure 6 based on horizontal line width results in an ordering of (largest to smallest) Crew, 3rd, 1st, and 2nd – which is not correct either. Using horizontal width is inviting, since the lines are centered around the middle of a level, facilitating this comparison.

In this paper, we propose a new approach to visualizing relationships between categorical variables in a parallel coordinates setting, called the *common angle* plot.

III. COMMON ANGLE PLOTS

A. Construction

Figure 7 shows a common angle plot of the same data as the hammock plot.

As in the previously discussed display types ribbons are drawn between categories with widths that are proportional to the number of records they represent.

In order to ensure that widths of all bands are comparable without any distortion, their slopes are artificially made the same in the following manner: assuming a vertical display as shown in figure 7, we modify the connecting bands between categories from a straight band to a combination of a vertical segment, a segment under a pre-specified angle θ, followed by another vertical segment. The pre-specified angle θ (between the line and the vertical band) is given as –at most– the angle of the longest connecting line between two categories of neighboring variables. This makes the width of ribbons comparable without being affected by the distortion, as all ribbons are sharing at least one segment under the same angle.

B. Origins

Historically, the common angle plot has its roots in Sankey diagrams (29). Sankey diagrams have their origin
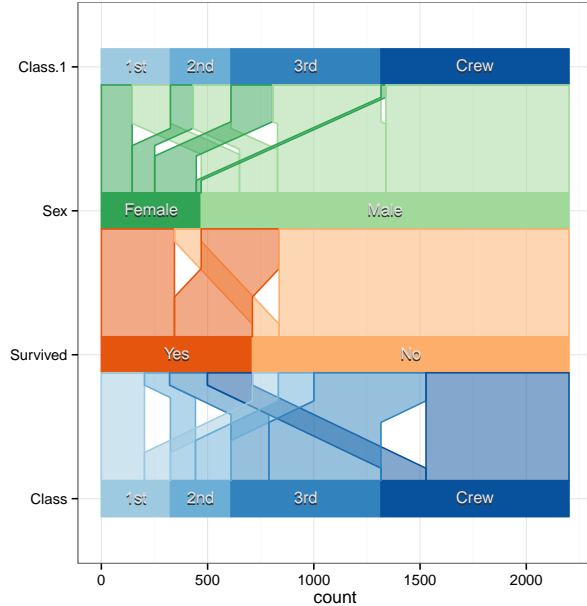

Fig. 7. Common angle plot of the Titanic data.

and main use of application in the area of industrial engineering. They are often used to show flows, particularly, in the framework of energy consumption and losses within systems. One of the most famous charts that can be classified as a Sankey diagram is Minard's map of Napoleon's march on Moscow (30). This map shows the size of the army as a flow chart from West to East and the subsequent retreat. Usually, Sankey diagrams do not show a geographic co-variate, though. Modern uses of Sankey diagrams are discussed in detail in (31); they can also be found in charts designed by Jen Christiansen, such as in (32) and (33).

A very recent example of related work can be found in (34); the chart is an interactive d3 display that shows political shifts of states and corresponding electoral votes in a combination of parallel coordinates of state-based votes overlaid by a Sankey-type plot to connect states across past elections.

C. Extensions

1) Hammock-adjusted common angle plots: As a logical next step in making ribbons comparable across all their parts and multiple variables, we can additionally adjust sloped parts of the ribbons using the hammock-adjustment described above. This leads to a representation as shown in figure 8. Note that ribbons no longer are required to fill the marginal bars. This choice was made to reduce the amount of overplotting resulting from the increase in width of the lines under an angle. The hammock-adjustment reintroduces the aspect ratio of the display as a crucial quantity, and with that the implied dependency on properties of the plotting device.
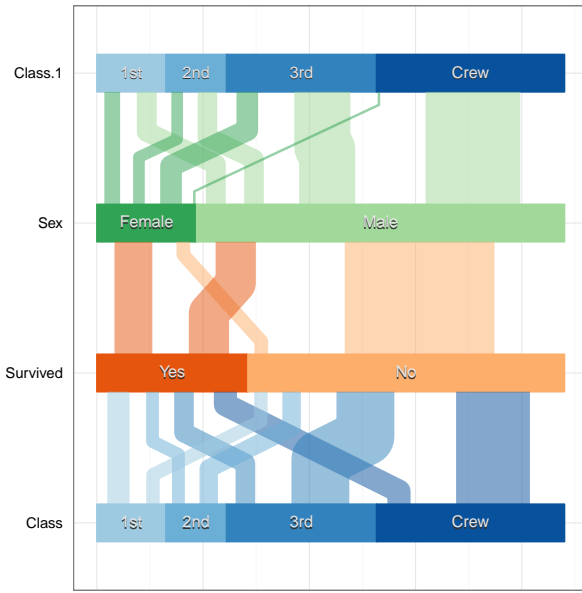

Fig. 8. Adjusted common angle plot of the Titanic data.

2) Hierarchies in common angle plots: In the original paper, parallel sets have been introduced to reflect a hierarchy of variables. In this paper, we have only used sets of two-dimensional plots to focus on the association between pairs of variables. However, it is possible, to show hierarchies in all of the types of displays by using a color coding introducing a hierarchical structure to the data.

Figure 9 shows a common-angle plot with a hierarchy: survivors of the disaster are marked in blue, non-survivors by orange. From top to bottom of the plot a hierarchy is drawn, considering first, survival, then gender, followed by age and finally, class membership. The coloring helps to keep track of survival status throughout the hierarchy, the layout in a common angle plot makes comparisons valid across all levels. This is of particular importance in hierarchical displays, as these displays, by definition, have a larger number of smaller groups than displays without a hierarchy exacerbating problems induced by the line width illusion.
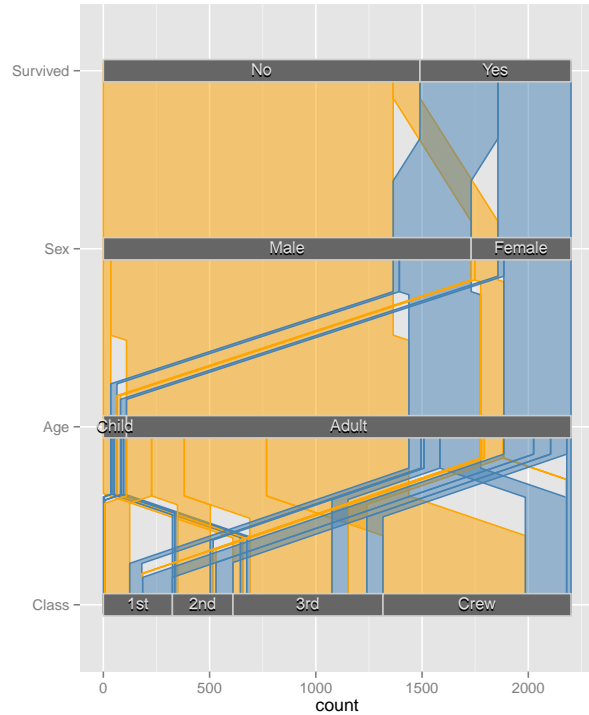

Fig. 9. Common angle plot of the Titanic data using a hierarchical structure in the variable (cf. to parallel sets chart in (24)).

IV. EVALUATION

A. Survey Design

To determine the effectiveness of the common angle plot, we conducted a user experiment in form of a survey asking participants to provide responses regarding the structure in two data sets with predominantly categorical variables.

All individuals were presented with the same six questions, in two blocks of three questions each. The first block of questions investigated structures in the Titanic data set, including a follow-up of the survival-class ordering issue discussed earlier. Participants were exposed to one of the three types of displays discussed: a parallel sets plot, the hammock plot of figure 6, or the common angle plot of figure 7.

The second block of questions regarded gene-pathway relationships in a dataset constructed from the UCSC Genome Browser (35) for human genome assembly.

Screening public databases of biochemical pathway and genetic sequence data to identify targets for further analysis is a common exploratory analysis task that benefits from visualization methods. One application is to determine if a drug that is known to affect a particular pathway may also affect function on a different pathway. Such information is useful in identifying new product markets or accurate warning labels. Another application is the discovery of co-localization of genes involved in multiple pathways; physical proximity is one indicator of evolutionary origin and may also provide clues regarding the mechanism that a pathway employs.

The pathways we focus on here, are involved in caffeine metabolism (hsa00232), steroid biosynthesis (hsa00100), and hsa00982, which is involved with drug metabolism via the cytochrome P450 superfamily. For simplicity, we also focus on just six different chromosomes instead of the regular 22. All data manipulation was performed using software R 2.15.1 (25) with Bioconductor package KEGG.db (36).
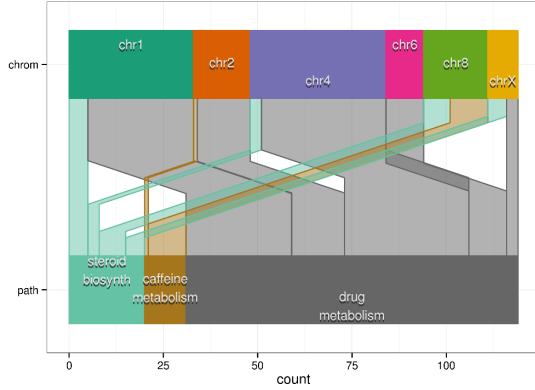

Fig. 10. Common angle plot of chromosome and pathway relationship. Line width corresponds to number of genes pathways and chromosomes have in common. Both drug metabolism (hsa00982) and caffeine metabolism (hsa00232) make use of genes on chromsome 2.

The pathway-chromosome relationship is shown in a common angle plot in figure 10. The top bar identifies chromosome location, the bottom pathways. Ribbon widths visualize the number of genes on a chromosome that are active in a particular pathway. A chromosome with ribbons connecting to multiple pathways is a physical target for drugs that possibly affect both paths. Gene CYP1A2 is involved in both caffeine metabolism and drug metabolism - clomipramine (an antidepressant) and caffeine both may act as a substrate for this gene. From a user perspective screening by the meta-location of chromosome reduces the question of co-localization of genes to one of searching for common elements from amongst 15 genes - reducing the search space by over 90% (to 15 from 193 total genes listed for the pathways).

Alternative displays are graph networks that usually focus on the display of nodes within a pathway, or circos displays (37), which generally emphasize the chromosome-chromosome interaction.

Common angle plots of gene-pathway interactions fit into Shneiderman's Visual Information Seeking mantra 'Overview first, zoom and filter, then details-on-demand' (38) as the first step that allows us to get an overview of the data and easily identify areas of potential interest for further investigation.

Participants were asked a second set of three questions (see appendix A for details) on this data set, but they were exposed to a type of display different from the type accompanying questions from block 1.

This resulted in six different survey administrations, that were answered by similar numbers of participants, as shown in Table I. The crossover design allows us to compare designs across different questions and data sets, while at the same time allowing to adjust for individuals' different skill sets.

Block 1	PS	CA	Ham	PS	CA	Ham
Block 2	CA	PS	PS	Ham	Ham	CA
#responses	8 (9)	6 (7)	8 (9)	6 (7)	10 (11)	8 (8)

TABLE I
OVERVIEW OF STUDY DESIGN AND PARTICIPATION NUMBERS. THE NUMBER IN PARENTHESIS INDICATES THE NUMBER OF PARTICIPANTS COMPLETING THE FIRST BLOCK, BUT NOT THE SECOND.

For the analysis, we first fit two linear mixed effects models, $M1$ and $M2$ – the first one gives a general overview of comparing the designs, the second one further highlights the performance of the designs in each question. We will then proceed to investigate some of the results in more detail.

B. Results and Analysis

1) Mixed effects models: Answers for each question on the survey were assessed for correctness, giving us a binary variable of 1s and 0s, which we used to investigate performance of different designs.

In a first model, $M1$, we are only interested in the overall difference in performance between designs. We can express this as a model of the form

$$[M1] \qquad g(P(y_i = 1)) = \mu + d_{j(i)} + u_{k(i)} + \varepsilon_i \qquad (2)$$

where $P(y_i = 1)$ is the probability that the ith response is correct, for $i = 1, ..., n$. $g(.)$ is the transformation (link function) of the response. Here, we make use of the logit link, i.e.

$$g(P(y_i = 1)) = \text{logit } P(y_i = 1) = \\ = \log P(y_i = 1) - \log P(y_i = 0).$$

$d_{j(i)}$ is the parameter measuring the effect of design j ($j = C, H, P$ for Common Angle, Hammock plot, and Parallel sets plot), $u_{k(i)}$ is the effect of participant's k individual skills in evaluating these plots, $k \in \{1, ..., 52\}$. The assumption is that skills are independent and normally distributed with an expected value of zero and a variance of σ_u^2. ε_i is, similarly to a regular linear model, assumed to be independently distributed according to a normal distribution with mean of zero and variance σ^2.

All models are fit in the lme4 package (39) within the software frame work of R 2.15.1 (25). Table II gives an overview of the model coefficients and their estimates. The effect of the common angle plot is used as a baseline, i.e all the effects shown are differences with respect to the performance of common angle plots. Both hammock plots and parallel sets have negative effects on the correctness of the response. This indicates a significantly worse performance of these designs than under the common angle plot.

Table III shows percentages of correctness for each design and each question. Bold numbers indicate significantly different (worse) performance of a design compared to the

common angle plot. Note that question A.2 overall had the worst performance, parallel sets exhibit significantly worse performance in four out of the six questions, hammock plots are inferior to the common angle plot in two out the six questions. Note that in the last question hammock plots have the best performance – but this is does not constitute a significant difference to the performance of the common angle plot. Also note, that the table shows raw percentages, i.e. these performance rates are not adjusted for skills of individuals.

	Estimate	Std. Error	z value	Pr($>$|z|)	
μ	1.72	0.24	7.04	0.00	***
d_C	0.00	–	–	–	
d_H	-0.86	0.32	-2.65	0.01	**
d_P	-1.31	0.32	-4.07	0.00	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

TABLE II

MODEL FIT FOR $M1$ MEASURING CORRECTNESS OF ANSWERS IN THE SURVEY. THE COMMON ANGLE PLOT IS USED AS BASELINE – BOTH HAMMOCK PLOTS AND, TO A LARGER DEGREE, PARALLEL SETS, SHOW A SIGNIFICANTLY WORSE PERFORMANCE.

	common	hammock	parallel
A1	86.0	76.5	81.2
A2	63.2	**5.9**	**12.5**
A3	73.7	**17.6**	**25.0**
B1	93.8	83.3	82.2
B2	68.8	68.8	**6.7**
B3	75.0	87.5	**53.3**

TABLE III

RAW PERCENTAGES OF CORRECTNESS OF RESPONSES FOR EACH QUESTION AND DESIGN. BOLD NUMBERS INDICATE SIGNIFICANT DIFFERENCE FROM COMMON ANGLE PLOT PERFORMANCE.

Model $M2$ extends the first model by including both effects for individual questions and the interaction effects with each design in the following form:

$$[M2] \quad g(P(y_i = 1)) =$$
$$\mu + d_{j(i)} + q_{q(i)} + p_{j(i),q(i)} + u_{k(i)} + \varepsilon_i \quad (3)$$

Here, q indicates the effects on performance for each question, where $q(i)$ describes one of the questions $\{A1, A2, A3, B1, B2, B3\}$. p is the parameter for the interaction effect of design and question, its index is a tuple consisting of a combination of a question and design.

Figure 11 gives an overview of the goodness of fit of model $M2$ – histograms of fitted values are drawn facetted by levels of the dependent variable. For correct responses fitted values are highly left skewed, with most values above 0.5. For wrong answers we see a symmetric, if not quite as clear-cut picture: fitted values are skewed right. Overall, Model $M2$ is a significant improvement over model $M1$ (a corresponding log-likelihood ratio test is significant at a level of $\ll 10^{-8}$). Table IV shows an overview of the parameters and their estimates for model $M2$. After adjusting for individuals' skills parallel sets
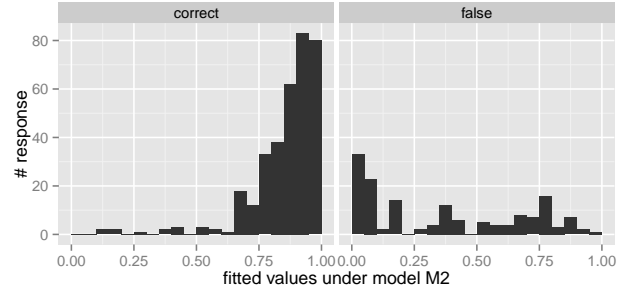

Fig. 11. Histograms of fitted values under model $M2$, facetted by actual performance of participants. On the left, correct responses are shown. The histogram of fitted values is skewed to the right, with most values above 0.5. The histogram on the right corresponds to wrong answers. There are fewer wrong answers, and they tend to have low fitted values, but there are more false positives among them than false negatives for correct answers.

	Estimate	Std. Error	z value	Pr($>$|z|)	
μ	2.68	0.61	4.35	0.00	***
Design					
d_C	0.00	–	–	–	
d_H	-1.06	0.83	-1.28	0.20	
d_P	-0.57	0.86	-0.66	0.51	
Questions					
q_{A1}	0.00	–	–	–	
q_{A2}	-1.85	0.74	-2.50	0.01	*
q_{A3}	-1.11	0.77	-1.44	0.15	
q_{B1}	0.55	0.96	0.58	0.56	
q_{B2}	-1.69	0.88	-1.91	0.06	.
q_{B3}	-1.31	0.92	-1.43	0.15	
Interaction					
$p_{H,A1}$	0.00	–	–	–	
$p_{P,A1}$	0.00	–	–	–	
$p_{H,A2}$	-3.29	1.55	-2.13	0.03	*
$p_{P,A2}$	-3.03	1.33	-2.28	0.02	*
$p_{H,A3}$	-2.61	1.17	-2.24	0.03	*
$p_{P,A3}$	-2.60	1.16	-2.23	0.03	*
$p_{H,B1}$	-0.26	1.21	-0.22	0.83	
$p_{P,B1}$	-0.81	1.24	-0.65	0.51	
$p_{H,B2}$	1.03	1.22	0.85	0.40	
$p_{P,B2}$	-3.48	1.63	-2.14	0.03	*
$p_{H,B3}$	1.97	1.37	1.44	0.15	
$p_{P,B3}$	-0.62	1.26	-0.50	0.62	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

TABLE IV

MODEL FIT FOR $M2$ MEASURING CORRECTNESS OF ANSWERS IN THE SURVEY, DETAILING PERFORMANCE OF DESIGNS ON EACH QUESTION. ALL DESIGN COMPARISONS ARE WITH RESPECT TO THE COMMON ANGLE PLOT.

performs significantly worse than common angle plots in three of the six questions.

Figure 12 shows an overview of the predicted skill for each participant under model $M2$. Skills are quite varied between -1.52 and 1.34, but a Kolmogorov-Smirnov test does not show significant deviation from a normal assumption (p-value = 0.089). On the scale of the dependent variable the range in individuals' skills translates to a $17.5 = e^{1.34-(-1.52)}$ fold increase in the probability of answering a question on the survey correctly between participants with the best skill set and the worst.

In the next section we will investigate the results from the

Fig. 12. Histogram of the predictions of subject-specific skills.

survey in further detail for some questions, and highlight the link to the line width illusion and its inverse.

2) Evidence for line width illusions: Question $A2$ asked participants to order class levels according to the number of survivors, fewest to highest. There are $4! = 24$ distinct orderings of the levels, corresponding to all permutations of length 4. Some orderings are closer to one another than other orderings. To quantify distance between each pair of permutations, the Cayley distance is defined as the minimal number of transpositions (i.e. swaps of two elements) necessary to transform one permutation into another.

This distance is used to define a graph on the space of all permutations: let all permutations be nodes of the graph; for each two permutations with a Cayley distance of one, we add an edge between these nodes to the graph. The resultant regular graph of degree six, i.e. each node is connected to six other nodes. Between any two nodes, the Cayley distance is then given as length of shortest connecting path shown on the graph. Figure 13 shows an overview of the permutation space together with an overview of the survey results. Each permutation is represent by one dot.

The colored dots on top of the graph correspond to the responses from the survey - the size of the dots is proportional to the number of observers choosing this particular ordering. It becomes obvious from the three graphs in figure 13 that the answers to different designs occupy quite different regions, while answers based on the same design are quite close – usually separated by only one edge.

The correct ordering, as well as the orderings assuming the line width illusion and its reverse are marked with symbols. Answers for the common angle plot are centered around the correct answer, while responses to parallel sets are cluster around the response corresponding to the line width illusion. Due to the small number of responses to the hammock design a clear clustering of the answers is not recognizable, but the answer for the inverse line width illusion is among the responses. Table V gives an overview of all responses to question $A.2$.

Common angle plots show the best performance in terms of correctness (66.7% on 18 responses), compared to a correctness of 11.8% for parallel sets plots on 16 responses, indicating a significantly better performance of common angle plots at a level of 0.0027, based on a Mantel-Haenszel test (the difference in performance to hammock plots is also significant at a level of 0.0006; but there is no significant difference in correctness between hammock plots and parallel sets.). While the intuitive assessment of

Order	CA	H	PS	
Crew, 1st, 3rd, 2nd		2		
2nd, 1st, 3rd, Crew		6		inverse illusion
Crew, 3rd, 1st, 2nd		1	1	
2nd, 3rd, 1st, Crew	1	7	1	
2nd, 3rd, Crew, 1st	12	1	2	correct
Crew, 3rd, 2nd, 1st	2			
3rd, 2nd, Crew, 1st	1			
1st, 2nd, 3rd, Crew	1			
1st, 3rd, Crew, 2nd	1		2	
Crew, 2nd, 3rd, 1st			6	line width illusion
1st, 3rd, 2nd, Crew			2	
2nd, Crew, 3rd, 1st			3	
Total	18	17	16	

TABLE V
RESPONSES TO QUESTION $A.2$: ORDER LEVELS OF CLASS BY THE NUMBER OF SURVIVORS (SMALLEST TO LARGEST).

lines by their width orthogonal to their direction is well known, it is surprising to see its strength: in this particular setting, it is strong enough to 'shrink' the horizontally widest line for 6 out of 16 participants by at least 44%, from 212 to below 118, and a further 3 participants perceived the shrinking to below 178, which constitutes a distortion of at least 16%.

Responses to question A.3 exhibit a similar pattern, see table VI.

Order	CA	H	PS	
c, b, a		2		
a, b, c	1	12		inverse illusion
a, c, b	13	3	4	correct
b, c, a	2	1		
c, a, b	2		9	line width illusion
b, a, c			2	
Total	18	17	16	

TABLE VI
RESPONSES TO QUESTION A.3: ORDER COMBINATIONS FROM SMALLEST TO LARGEST, WHERE 'A' IS FIRST CLASS FEMALE, 'B' ARE MALE SURVIVORS, AND 'C' ARE CREW SURVIVORS.

Qu	Design	Correct	Incorrect	No Answer	Reason
1	common	14	2	0	
	hammock	15	0	1	
	parallel	7	7	0	line width illusion
2	common	16	0	0	
	hammock	9	7	0	inverse illusion
	parallel	14	0	0	
3	common	15	1	0	
	hammock	16	0	0	
	parallel	14	0	0	

TABLE VII
RESPONSES TO QUESTION SET B.1

Table VII shows a summary to the three questions of $B.1$ with possible answers "agree", "disagree", and "don't know". The first two questions are examples, where the line width illusion, and its reverse will lead to answers that differ from the correct answer. Parallel sets are susceptible to the line width illusion, while hammock plots suffer from
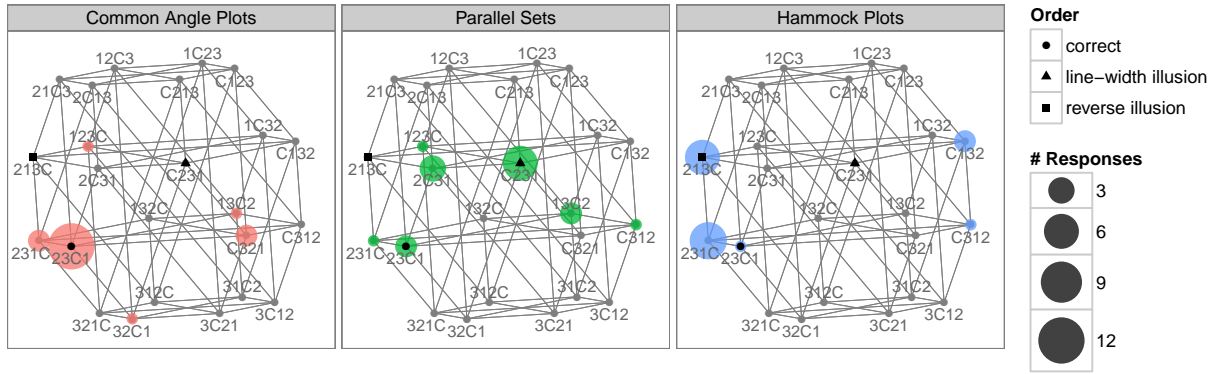
Fig. 13.   Answers to question A.2 – each node corresponds to a single ordering of the levels in variable 'Class'. Lines are drawn between orderings that are only one swap of levels apart. The colored dots show responses from the survey, their sizes depend on the number of responses for each ordering.

the reverse. The data shows that in about 50% of the responses we can see this difference. 7 out of 14 answers for the parallel sets plots in question 1 deviate from the correct answer, and 7 out of 16 responses corresponding to hammock plots in question 2 show the wrong answer.

3) Opinion on common angle plots: Answers to the question of 'which chart did you like better?' are shown in table VIII. There is a pretty clear endorsement in favor of common angle plots versus the other two types of displays. Asked for a reason for their preference over parallel sets, 8 participants cited a facilitated comparison of width, area or "size", 3 saw the common angle plots as preferable, and the remaining answers related to a 'more logical' structure or generally 'easier' choice. Reasons for not preferring common angle plots boiled down to a preference for straight lines. This purely aesthetic preference is deeply rooted and in our opinion the biggest challenge for common angle plots.

Which chart did you like better?			Chart 1	Chart 2
PS	vs	**CA**	2	**6**
CA	vs	PS	**4**	2
H	vs	**CA**	3	**5**
CA	vs	H	**8**	2
H	vs	PS	3	5
PS	vs	H	1	5

TABLE VIII
PREFERENCES FOR FIRST OR SECOND CHART ACROSS ALL SIX
COMBINATIONS OF QUESTIONS AND CHART TYPES.

V. IMPLEMENTATION

All of the discussed variants of parallel coordinates for categorical data are implemented as part of the package `ggparallel` based on the `ggplot2` (40) plotting framework in the software R 2.15.1 (25). The `ggparallel` package is freely available from CRAN (http://www. r-project.org/). The colors for the plots have been chosen using color schemes from the ColorBrewer project (41) , as implemented in the R package `RColorBrewer` (42) .

VI. CONCLUSION

We have been investigating different types of plots used for visualizing structures between categorical variables, specifically parallel sets and hammock plots. We have presented strong evidence that hammock plots and parallel sets suffer from a type of Müller-Lyer illusion, which we have coined the line width illusion and its reverse. We have introduced common angle plots as a means to avoid the line width illusions. We have found strong evidence that the common angle plot is not affected by these line width illusions with significantly better performance in a user study.

We have presented overwhelming evidence in the 'case against straight lines', particularly when line width is encoding quantitative information. Although straight lines are a simple and mathematically perfect solution for representing connections between levels of two variables, they are far from optimal for the peruser of the chart. Parallel sets are subject to the line width illusion, and the adjustment seen in hammock plots are prone to the reverse of the line width illusion. The line width illusion is a symptom of the tension between mathematical and intuitive approaches to elucidating plot encoding. While mathematics tells us that we need to assess line widths in a vertical or horizontal fashion to get a correct evaluation of the encoded information, there is also a strong innate tendency to evaluate widths of lines orthogonally to their direction. Adjusting for the line width illusion by changing the width corresponding to the slope of the line segment such as in hammock plots makes either end of the segments susceptible to the reverse line width illusion. We found strong support in the survey we conducted for the existence of both the line width illusion in parallel sets and the reverse in hammock plots. By giving up on the paradigm of straight lines for drawing connections between levels in the common angle plot, we are able to circumvent both illusions, and present an illusion-free alternative that encodes the same information as the widely used parallel sets or the hammock plots.

APPENDIX

Survey

The survey was created using the Qualtrics Labs, Inc software (www.qualtrics.com). At the survey start, participants were presented a brief tutorial regarding the different plot types. The tutorial can be found at http://mariev.net/tutorial.html

The survey consists of two blocks of questions each pertaining to one data set (titanic data or chromosome-pathway interactions). Each block was presented with a single plot to use as reference when responding. Two different plot types were shown for the two different blocks, yielding a total of six unique orderings of plot types (CA-PS, PS-CA, CA-H, H-CA, H-PS, and PS-H).

Participants were randomly assigned to one of these six combinations, each of approximately equal size. This study design structure was imposed, in part, to encourage participation by reducing the amount of time for survey completion. Completion of all survey questions was anticipated to take 10 - 15 minutes. No personally identifiable information was collected, nor was any compensation offered.

The questions pertaining to the titanic dataset were:

A.1 *Agree, Disagree or Don't Know/Can't Determine with the following statements:*
  – There were an approximately equal number of Male and Female Survivors
  – The group with largest number of travelers was Female Survivors
  – There were more Male Non-Survivors than number of males in First and Second Class Combined

A.2 *Order the categories of Class by number Survived, fewest to most.*
  – 1st
  – 2nd
  – 3rd
  – Crew

A.3 *Order the following groups by number, fewest to most*
  – 1st Class female passengers
  – Male Survivors
  – Crew Survivors

The questions pertaining to the gene dataset were:

B.1 *Agree, Disagree or Don't Know/Can't Determine with the following statements:*
  – There are about the same number of genes in the group "steroid biosynthesis:chromosome 1" as in the group "caffeine metabolism: chromosome 8"
  – The group with the greatest number of genes is "drug metabolism:chromosome 4"
  – there are more genes involved in the group "drug metabolism: chromosome 1" than all genes involved in the caffeine metabolism pathway

B.2 *Order the following chromosomes by number of genes involved in steroid biosynthesis pathway, fewest to most.*
  – chromosome 1
  – chromosome 4
  – chromosome 8
  – chromosome X

B.3 *Order the following chromosomes by number of genes involved, fewest to most.*
  – steroid biosynthesis :: chromosome X
  – steroid biosynthesis :: chromosome 4
  – drug metabolism :: chromosome X

Participants' demographics

All students, staff and faculty from Iowa State University programs in Statistics, Bioinformatics and Computational Biology and Human Computer Interaction were invited to participate by email. 93 individuals accessed the survey; 86 participants gave consent, 15 of those dropped out right after, 20 went to the training site and did not return. Out of the remaining 51 participants, 46 individuals submitted responses for all questions and 5 gave responses to the first block of questions.

Participants used their own personal computing devices to access the survey, a majority of participants used Intel Mac OS X (versions ranging from 10.6.8 to 10.8.2), with the next most common operating system WoW64 (Windows 32-bit or Windows 64-bit). The preferred choice of browser was Firefox, followed by Chrome. For two participants, the Qualtrics survey software was unable to capture operating system or browser information.

REFERENCES

[1] E. Tufte, *The Visual Display of Quantitative Information*, 2nd ed. USA: Graphics Press, 1991.

[2] M. d'Ocagne, "Coordonnées parallèles et axiales : Méthode de transformation géométrique et procédé nouveau de calcul graphique déduits de la considération des coordonnées parallèles," Gauthier-Villars, Paris, 1885.

[3] A. Inselberg, "The plane with parallel coordinates," *The Visual Computer*, vol. 1, no. 2, pp. 69–91, 1985.

[4] E. J. Wegman, "Hyperdimensional data analysis using parallel coordinates," *Journal of the American Statistical Association*, vol. 85, no. 411, pp. pp. 664–675, 1990. [Online]. Available: http://www.jstor.org/stable/2290001

[5] R. Kosara, F. Bendix, and H. Hauser, "Parallel sets: Interactive exploration and visual analysis of categorical data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 4, pp. 558–568, Jul. 2006. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2006.76

[6] F. Bendix, R. Kosara, and H. Hauser, "Parallel Sets: Visual analysis of categorical data," in *IEEE Symposium on Information Visualization*, October 2005.

[7] J. LeBlanc, M. O. Ward, and N. Wittels, "Exploring n-dimensional databases," in *Proceedings Visualization*. IEEE CS PRess, 1990, pp. 230–237.

[8] M. Spenke and C. Beiken, "Visualization of trees as highly compressed tables with InfoZoom," in *Proceedings IEEE Information Visualization*, 2003, pp. 122–123.

[9] M. Schonlau, "Visualizing categorical data arising in the health sciences using hammock plots," in *Proceedings of the Section on Statistical Graphics*, RAND Corporation. American Statistical Association, 2003.

[10] J. Hartigan and B. Kleiner, "Mosaics for contingency tables," in *Proceedings Symposium on the Interface*, 1982.

[11] M. Friendly, "Visualizing categorical data: Data, stories and pictures," in *SAS User Group Conference*, 1992.

[12] H. Hofmann, "Exploring categorical data: Interactive mosaic plots," *Metrika*, 2000.

[13] M. Theus, H. Hofmann, B. Siegl, and A. Unwin, *New Techniques and Technologies for Statistics II*. IOS Press Amsterdam, 1997, ch. MANET: Extensions to interactive statistical graphics for missing values.

[14] W. S. Cleveland and R. McGill, "Graphical perception: Theory, experimentation, and application to the development of graphical methods," *Journal of the American Statistical Association*, vol. 79, no. 387, pp. pp. 531–554, 1984. [Online]. Available: http://www.jstor.org/stable/2288400

[15] W. Playfair, *Commercial and Political Atlas*, London, 1786.

[16] W. Playfair, H. Wainer, and I. Spence, *Playfair's Commercial and Political Atlas and Statistical Breviary*. Cambridge University Press, 2005.

[17] H. Wainer, *Visual Revelations*. Psychology Press, 2000.

[18] N. Robbins, *Creating More Effective Graphs*. Wiley, 2005.

[19] R. H. Day and E. J. Stecher, "Sine of an illusion," *Perception*, vol. 20, pp. 49–55, 1991.

[20] M. Blastland, "Go figure: How to understand risk in 13 clicks," March 11 2009. [Online]. Available: http://news.bbc.co.uk/2/hi/uk{_} news/magazine/7937382.stm

[21] R. Kosara, "Redesigning parallel sets," Visweek 2008 workshop, 2008. [Online]. Available: http://www.stonesc.com/Vis08_Workshop/image%20array.htm

[22] ——. Eagereyes. [Online]. Available: http://eagereyes.org/parallel-sets

[23] M. Bostock, V. Ogievetsky, and J. Heer, "D3: Data-driven documents," *IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis)*, 2011. [Online]. Available: http://vis.stanford.edu/papers/d3

[24] J. Davies. Parallel sets. [Online]. Available: http://www.jasondavies.com/parallel-sets/

[25] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012, ISBN 3-900051-07-0. [Online]. Available: http://www.R-project.org/

[26] R. J. Dawson, "The 'unusual episode' data revisited." *Journal of Statistics Education*, vol. 3, 1995. [Online]. Available: http://www.amstat.org/publications/jse/v3n3/datasets.dawson.html

[27] J. Heer and M. Bostock, "Crowdsourcing graphical perception: Using mechanical turk to assess visualization design," in *CHI 2010: Visualization*, 2010.

[28] N. Kong, J. Heer, and M. Agrawala, "Perceptual guidelines for creating rectangular treemaps," *Transactions on Visualization and Computer Graphics*, 2010.

[29] H. R. Sankey, "Introductory note on the thermal efficiency of steam-engines," Report of the Committee appointed on the 31st March, 1896, to Consider and Report to the Council upon the Subject of the Definition of a Standard or Standards of Thermal Efficiency for Steam-Engines: with an Introductory Note. Minutes of Proceedings of The Institution of Civil Engineers, pp. 278–283, 1898.

[30] C. Minard, "Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813," 1869.

[31] M. Schmidt, "The Sankey diagram in energy and material flow management," *Journal of Industrial Ecology*, no. 12, pp. 173–185, 2008.

[32] L. L. National Laboratory and the U.S. Department of Energy, "The efficiency solution," *Scientific American*, June 2009.

[33] M. Fischetti, "Water in, water out," *Scientific American*, June 2012.

[34] S. Carter and M. Bostock. (Oct 15 2012) Over the decades, how states have shifted. [Online]. Available: http://www.nytimes.com/interactive/2012/10/15/us/politics/swing-history.html

[35] W. Kent, C. Sugnet, T. Furey, K. Roskin, T. Pringle, A. Zahler, and D. Haussler, "The human genome browser at UCSC," *Genome Research*, vol. 12, no. 6, pp. 996–1006, June 2002.

[36] M. Carlson, S. Falcon, H. Pages, and N. Li, *KEGG.db: A set of annotation maps for KEGG*.

[37] M. I. Krzywinski, J. E. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, and M. A. Marra, "Circos: An information aesthetic for comparative genomics," *Genome Research*, 2009. [Online]. Available: http://genome.cshlp.org/content/early/2009/06/15/gr.092759.109.abstract

[38] B. Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations," in *Proceedings of the IEEE Symposium on Visual Languages*. IEEE Computer Society Press, 1996, pp. 336–343. [Online]. Available: http://citeseer.ist.psu.edu/409647.html

[39] D. Bates, M. Maechler, and B. Bolker, *lme4: Linear mixed-effects models using S4 classes*, 2012, R package version 0.999999-0. [Online]. Available: http://CRAN.R-project.org/package=lme4

[40] H. Wickham, *ggplot2: elegant graphics for data*

analysis. Springer New York, 2009. [Online]. Available: http://had.co.nz/ggplot2/book

[41] M. A. Harrower and C. A. Brewer, "ColorBrewer.org: An Online Tool for Selecting Color Schemes for Maps," *The Cartographic Journal*, vol. 40, no. 1, pp. 27–37, 2003.

[42] E. Neuwirth, *RColorBrewer: ColorBrewer palettes*, 2011, R package version 1.0-5. [Online]. Available: http://CRAN.R-project.org/package=RColorBrewer

Heike Hofmann is associate professor of Statistics at Iowa State University. She is a member of the inter-disciplinary Human Computer Interaction program at Iowa State and serves as faculty member for the Bioinformatics and Computational Biology program. Her research interests are the visual exploration of high-dimensional and large data. Her research group has been the recipient of several awards for their data detective and visualization skills.

Marie Vendettuoli is a graduate student at Iowa State University, with majors in Human Computer Interaction and Bioinfomatics & Computational Biology. She is a past IGERT Fellow and currently interns with the Statistics section at USDA. Her research interests involve the use of interactive visualizations in exploratory data analysis.