Publication V

Mikko Berg and Ilpo Kojo. 2012. Integrating complex information with object displays: psychophysical evaluation of outlines. Behaviour & Information Technology, volume 31, number 2, pages 155-169.

© 2012 Taylor & Francis

Reprinted by permission of Taylor & Francis.

Behaviour & Information Technology
Vol. 31, No. 2, February 2012, 155–169

Integrating complex information with object displays: psychophysical evaluation of outlines

Mikko Berg* and Ilpo Kojo

Helsinki School of Economics, Center for Knowledge and Innovation Research, PO Box 1210, 00101 Helsinki, Finland

(*Received 5 November 2008; final version received 7 December 2009*)

The interactive use of visual interface tools has diversified the use of visualisations. This article reviews the relevant aspects of interaction and challenges the sufficiency of traditional evaluation criteria developed for static graphs. Traditionally, the problem for statisticians has been to maintain perceptual discriminability of details, when quantities of data increase. Currently, however, even non-professional users need to integrate qualitatively different kinds of information. The review of task requirements indicates the use of a visual outline: (1) visual tools can facilitate parallel separation of individual data entities and integration of their features and (2) more focused comparisons require visual memory due to eye movements. The article reports psychophysical experiments that measure performance accuracy and response latency conditioned by the above task requirements. The impact of shape and colour on performance interacted with display times; the times were shorter (100 ms) or longer (1 s) than the duration of typical gaze fixation. The features of graphs in the experiments were derived from a popular internet service. Thus, we describe methods for evaluating visual components of real services and provide general guidelines for visual design of human–computer interaction.

Keywords: visualisation; psychophysics; multi-dimensional; exploration; perception; cognition

1. Introduction

The movement of one's head slightly is often the most effective way to experience three-dimensional spatial relations of objects in the environment. In fact, this reaction is so natural that some children try to move their heads to look behind a tree in a computer golf game. Unfortunately, visual exploration of objects' features or their relations in visual interfaces is not yet so intuitive. However, the example highlights the significance of interaction and possibilities for improving it in the visual modality.

This article describes the challenges involved in creating a visual interface with which non-professional users can explore multi-feature information. Scientists and engineers are not the only people who are often confronted with large amounts of complex data. Many social and economic decisions are related to similar challenges, such as processing complicated information in political elections or product descriptions. In addition, even engineers might abandon the task of visualisation of multi-dimensional data because they are unable to evaluate the statistical requirements of the underlying algorithms (Laine 2003). Our solution was to design simple complementary interactive graphs to explore the details of the results.

This article explains how to select a suitable graph for the given purpose and evaluate the effectiveness of alternative graphs. Temporal requirements are emphasised for interactive graphs as compared to traditional static representations. The new criteria for evaluation are discussed and compared to previous studies.

2. Needs for visual exploration

2.1. *Complexity in everyday decisions*

To navigate the menu items of an operating system is convenient, if there is an obvious hierarchy. However, large amounts of data can cause problems and the user might not understand the underlying structure. A common task, such as online shopping, can become complex when the number of relevant attributes of the product increases. Indeed, customers of IKEA made better complex (but not simple) decisions when relying on intuition rather than through conscious deliberation (Dijksterhuis *et al*. 2006, Shanks 2006). Comparably, analysing reasons (Reber 1976, Berry and Broadbent 1988, Levine *et al*. 1996) or verbalising visual mental representations (Schooler and Engstler-Schooler 1990) interfered with intuitive decisions and learning. The most likely cause is the capacity limit for the number of consciously processed items (four items in Luck and Vogel (1997), Cowan (2000), but see Cowan and Morey (2006), Bays and Husain (2008) for differences). Fortunately, visual tools can be used to surpass these limitations (e.g. by visual clustering)

*Corresponding author. Email: mikko.j.berg@helsinki.fi

ISSN 0144-929X print/ISSN 1362-3001 online
© 2012 Taylor & Francis
http://dx.doi.org/10.1080/01449290903544660
http://www.tandfonline.com

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

because the cognitive processes at stake depend on the form of visual representations (Kleinmuntz and Schkade 1993, Zhang 1997).

Surprisingly, the task of shopping shares objectives with visualisation of scientific information (Card *et al.* 1999) which attempts to improve the effectiveness of expert data mining. The term information visualisation initially referred to the goal of '. . . 2D and 3D animation to explore information and its structure' (Robertson *et al.* 1989, p. 10). In the model of that article, users do not merely receive the available information; they also manipulate visual objects, in the spirit of Neisser's iterative perceptual cycle (Neisser 1976). Furthermore, the effectiveness of visualisations depends on the selection of graphical features to encode the information (graphical perception theory; Cleveland and McGill 1984, Cleveland and McGill 1985). Thus, psychophysics that measures the perception of graphical features can reveal the effectiveness of visualisation and predict the quality of subsequent decisions.

2.2. *Benefits of visual modality*

Visualisations can overcome many cognitive limitations of the user. They are part of a distributed socio-technical system (Hutchins 1995a, b) and serve as external cognitive artefacts (Larkin and Simon 1987, Hutchins 1995a, b, Scaife and Rogers 1996, Zhang 1997, Cox 1999) that aid in mental processing. First, used visualisations should be compatible with the mental processes at stake. Otherwise, a mental re-representation of the problem is needed. For instance, compare the mental effort of calculations LXVIII $\times$ X and 68 $\times$ 10 (Zhang and Norman 1995). The difference is due to computational offloading when visualisation explicitly represents problem states (external memory).

Second, a graphical medium or a visual tool can be used to restrict possible misinterpretations through inherent constraints (Hutchins 1995a, Stenning and Oberlander 1995, Scaife and Rogers 1996). Notice that the forms of representations differentially limit the perceived quantities in the example above, by the lengths of simple strings of digits (also explicitly in stem-and-leaf graphs in Tukey (1977)). Third, visual comparisons need not operate on single numbers, but spatial segregation and clustering can reduce the number of mentally operated items (Pomerantz 1981, Treisman 1982, Vecera and Farah 1994, Ariely 2001). Furthermore, instantly grasping the number of units and noticing potentially missing data is beneficial (Schwartz 1971).

To summarise, the general motivation for using visual artefacts is reliance on perceptual inferences rather than limited capacity for symbolic inferences (see Bauer and Johnson-Laird (1993), Stenning and Oberlander (1995) for analogical examples in logical problem solving). For instance, dual-task experiments suggest that logical problem solving is independent of visuo-spatial working memory (Gilhooly *et al.* 1993). The benefits are most apparent when self-produced drawings make possible otherwise impossible mental operations (Chambers and Reisberg 1985).

2.3. *The benefits of computation*

In addition to visual modality, data mining algorithms of interactive visualisations can simplify the data. As a result, the complexity takes the form of multi-dimensionality. Simplification means that the dimensionality is reduced, for instance, by self-organising feature maps (SOM; Kohonen 1982, 2001) or multi-dimensional scaling (MDS; Shepard 1962). It has been argued that human cognition also reduces complex perceptual and semantic information into simple low dimensional conceptual spaces (Gärdenfors 2000).

So far, these methods have been used only for scientific purposes. However, we have developed SOM-based decision aids for non-experts in political elections (Berg *et al.* 2006).

From the user's point of view, multi-dimensional visualisation is a simplifying decision aid that must to be trusted (Shneiderman 2002). Unfortunately, users are not able to evaluate the assumptions of the underlying simplifications and therefore do not trust decision aids (Kleinmuntz 1990). These algorithms might lose meaningful information in the reduction process, so it is only natural to question their validity.

Our solution to these problems was to use simple diagrams. We identified a multi-dimensional everyday decision task of using internet candidate selector before a political election. Their popularity in Finland has been exceptional (Carlson and Strandberg 2005, Moring and Mykkänen 2005) and our design had nearly 100,000 users (Berg *et al.* 2006). The need for a new type of service arises from the quantity of available data: there were hundreds of candidates, each of whom answered a 20- or 25-item questionnaire designed by political journalists. The number of candidates and questionnaire items is too large for the typical capacity of human working memory. Thus, the traditional approach presents the best-matching candidates only after sequential presentation of the questions. In contrast, our design (Berg *et al.* (2006), but see also Kontkanen *et al.* (2000), Kaipainen *et al.* (2001)) included the between-candidate comparisons as a multi-dimensional data visualisation task. The exploration of SOM representation was complemented by sector diagrams (Figure 1), because these simple

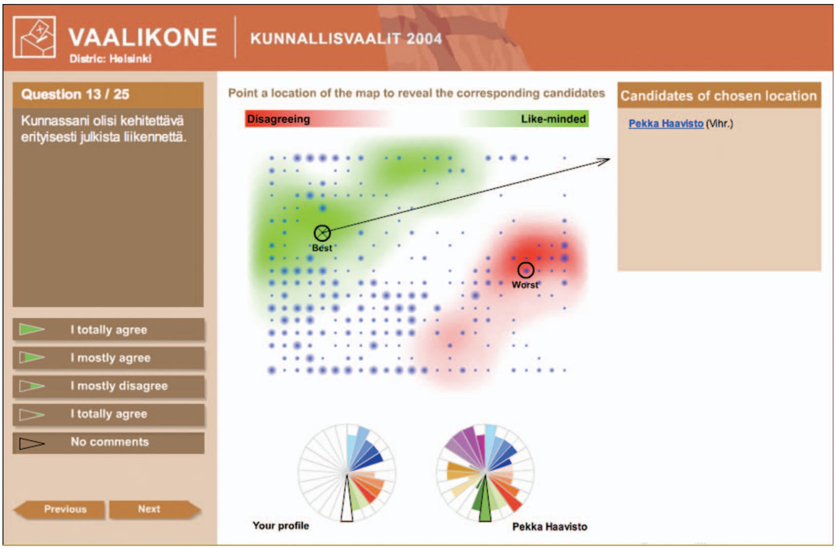

Figure 1. The candidate selector application consisted of four parts: particular item from questionnaire (on the left), the coloured SOM grid for presenting (in the middle) and selecting the candidates (on the left), and two sector graphs, one for the user's completed answers and one for the selected candidate.

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

diagrams provide, in general, an independent source for evaluating the correctness of the resulting clusters. The evaluation process itself is explorative process.

3. Perception of global patterns

3.1. *Importance of the overview*

A gold standard suggests several levels of detail in multi-dimensional data visualisation (Tufte 1983). For Bertin (1981), the sole problem of multi-dimensionality in visualisations is due to subdivisions that degrade overall vision. Moreover, *details-on-demand* principle states that overview should be presented before the details (e.g. Shneiderman 1996). The experiments on visual gist support this: semantic categories can be constructed in as little as in 100 ms, even though the picture would not be recognised little later (Potter 1976, 1993, Potter *et al*. 2002; see Spence (2002) for visualisation applications). Furthermore, the gist of a scene (Biederman 1972, Biederman *et al*. 1973, 1974) and semantics (Biederman 1982) affect object recognition on the same timescale. Unfortunately, only very few graphs are effective for both global and local applications (Simkin and Hastie 1987, Carswell *et al*. 1993). Specifically, Kosslyn's (1985, 1989) guidelines select line graphs for trends and interactions, and bar graphs for individual point values. However, both are often needed and perception begins at the overview.

We argue that the overview should be emphasised in the evaluation as well. The inventor of scientific visualisation, William Playfair (1786), makes the same point: 'Men of great rank, or active business, can only pay attention to general outlines . . . and it is hoped that with assistance of these charts, such information will be got, without fatigue and trouble of studying the particulars of which it is composed'. Moreover, educational systems have been criticised similarly. Observed difficulties in perceiving global patterns were attributed to disproportionate emphasis of school curricula on local interpretation of details (see also Guthrie (1988), Leinhardt *et al*. (1990)). The difficulties are emphasised by modern computer-based graphing tools that, unlike traditional numerical and algebraic approaches, require interpretation skills and the ability to generalise. A major problem is that students can often solve graphing or function problems but are unable to access scientific information from similar graphs.

3.2. *From discriminability to spatial integration*

Most visualisation psychophysics focuses on locating or discriminating details (Wainer and Thissen 1988), even though integration is more essential, difficult and time consuming (Vernon 1950, Guthrie 1988, Guthrie *et al*. 1993, Gattis and Holyoak 1996; eye

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

movement patterns in Carpenter and Shah (1998)). Especially, spatial separation imposes a cognitive load and disturbs learning (Sweller *et al.* 1990).

The focus on global patterns and integration across dimensions gives rise to different psychophysical criteria for evaluating graphs. Past experiments have typically measured latencies and performance accuracies[1] for individual data points represented in variants of bar and pie graphs. For instance, when two values were compared, the use of a simple bar chart (position among common scale) resulted in the most accurate decisions, divided bar charts (length) were the second most accurate, and pie charts (angle) were the least accurate (Cleveland and McGill 1984, 1985, Simkin and Hastie 1987). However, pie charts (and simple bar charts; Simkin and Hastie 1987) outperformed the divided bar chart when the data point was evaluated as a proportion of the whole (Eells 1926). Additional survey data indicated that most respondents spontaneously used absolute length judgments with bar graphs and proportional comparisons with pie charts (Simkin and Hastie 1987). Thus, the degree of effectiveness depends on the subtask emphasised.

Although most of the experiments have measured the distinctiveness of details (Cleveland and McGill 1984, 1985, Zacks *et al.* 1998), the same authors have admitted that 'the primary purpose of visualisations is not to convey numbers with as many decimal places as possible' (Cleveland and McGill 1984, p. 535). Furthermore, all the above experiments (but not ours) have compared the data values in isolation, thus neglecting the very strong contextual influence of neighbouring bars (Siegrist 1996, Zacks *et al.* 1998). Using just two bars is not enough.

Object displays have been specially designed for spatial integration (Carswell and Wickens 1987, Wickens and Hollands 1999, pp. 125–129, elsewhere glyphs; Ward 2002, or configurational displays; Bennet *et al.* 2000, Bennet and Walters 2001) as the data dimensions are represented by different visual features of a single object. First, the integration of information from different features of a single object is more accurate than from multiple objects (Lappin 1967). For instance, the use of a single triangle was superior to three different bars in an integration task, but not in other tasks (Casey and Wickens 1986, Carswell and Wickens 1987). Second, performing two different tasks simultaneously on one object rather than on two different objects is beneficial (Kramer *et al.* 1985) even if there is no difference in location when the objects are superimposed (Duncan 1984). The difference is due to so-called 'object-based attention', as opposed to 'spatial attention' (Vecera and Farah 1994). To summarise, the most important visual task is integration of relevant details. Appropriate spatial distances and object identification ease this task.

4. Benchmarking alternative graphs

4.1. *Tracking individuals across multiple dimensions*

Several simple graph types have been used for visual integration of high-dimensional data. First, a matrix using bivariate (2-D) scatterplots can be formed for each pair of dimensions (e.g. Chambers *et al.* 1983, Jacoby 1998). The idea is simple, but the result is extremely complex. Even when the redundant halves of the scatterplots are removed (draftman's display), our 20-dimensional data yield 190 plots with several hundred points (individuals) each. In addition, to gain coherent idea of a single individual one must track the correspondence of points between multiple plots. Parallel coordinate graph is another method to reduce the number of plots by sacrificing the correlations between individual dimensions. As a result, the x-axis represents the dimensions and the y-axis represents the data values aligned with the corresponding dimension. In parallel coordinates (Inselberg and Dimsdale 1990, Inselberg 1999) the values on each dimension for one individual are connected with a line (see Figure 2 for our data). For instance, in our election candidate data, finding a public figure for a reference point might be helpful. Again, individuals are difficult to distinguish. In another words, it is often important to find entities with commonness on several dimensions.

The individuals can be differentiated using profile (i.e. polygon) or star (i.e. circular line graph) graphs (Figure 3B and 3C, respectively). A profile graph is created from a parallel coordinate graph by separating into different graphs (Bertin 1967, Chambers *et al.* 1983, p. 163, Jacoby 1998, p. 20). Bertin (1967, p. 245) suggested the use of profile graphs for open-ended classification that are important for multi-dimensional visualisations. Interestingly, star graphs create more dramatic and memorable shapes with their radial representations (Chambers *et al.* 1983). The short-term visual memory (Luck and Vogel 1997, Cowan and Morey 2006, Bays and Husain 2008) is important for comparing large numbers of graphs with a sequence of eye movements.

4.2. *Avoiding arbitrary linking of dimensions*

Unfortunately, parallel coordinate, profile, and star graphs pose arbitrary links between otherwise independent dimensions. For one, empirical researchers have been encouraged to connect data points by lines only when the points have a numerical connection and to use different bars for different nominal categories (Cozby 2004, p. 219). Second, line graphs predispose

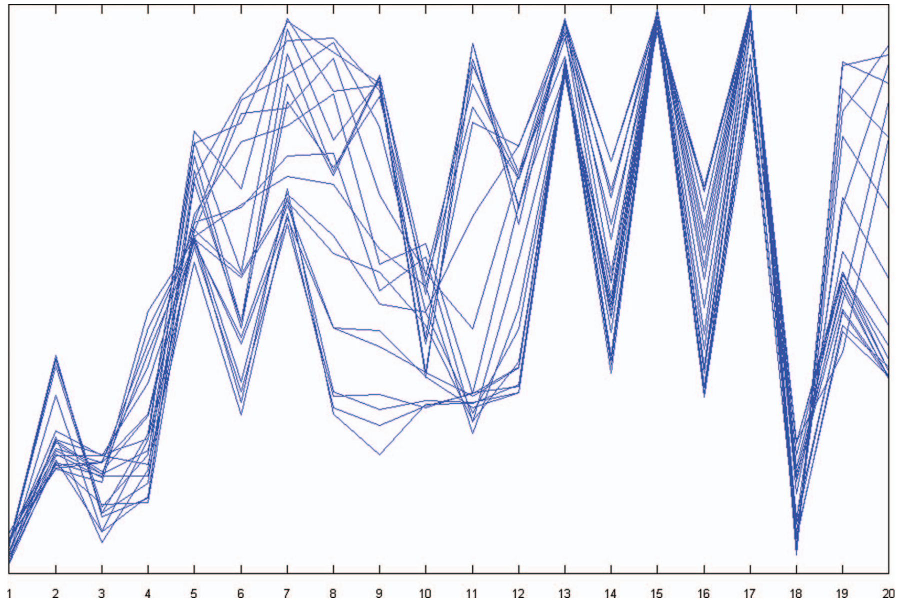

Figure 2. A parallel coordinate graph represents all 20 dimensions in very compact form and reveals the dimensions most and least (3 and 10) variance.

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

for a particular trend, while bar graphs allow considering more relations of discrete values (Shah *et al.* 1999, Zacks and Tversky 1999). Although neighbouring bars affect the perception of bar length, as mentioned earlier, the influence is less than with lines. In fact, it was the lack of continuous time series that led William Playfair (1786) to invent individual bars (Beniger and Robyn 1978).

Third, the Gestalt principle of connectedness (Wertheimer 1938) provides faulty signals of relatedness. Fourth, arbitrary connections and biases are likely reasons that the famous Chernoff faces (Chernoff 1973) are currently not popular in multi-dimensional visualisation (Ward 2002; for exceptions see Hahn *et al.* (1983)). Finally, lines connecting points create (Poggendorff's) geometric illusions (Poulton 1985). For some data, the connections are not arbitrary when the dimensions form meaningful groups; but note that the order of dimensions significantly influences the perception of outlines. In general, arbitrary coupling of data points distort perception of multi-dimensional data.

4.3. Selecting a bar or sector graph

For the above reasons, bar graphs (histogram plots in Jacoby (1998) and less familiar sector graphs (Figure 3A and 3D, respectively) are more objective alternatives to profile and star graphs. They differentiate individuals and do not create arbitrary links between dimensions. In sector graphs, data values on different dimensions are represented by lengths of sectors (polar construct in Bertin (1967) or Florence Nightingale's rose diagram after the inventor). As with pie charts, a user is expected to be primed to evaluate all quantities holistically. In fact, the following experiments compare radial length as a graphical code to simple bar graphs. Furthermore, sector graphs may inherit the benefits of the radial glyphs that outperformed scatter plots in complex data analysis (Anderson 1957). The sector graphs were preferred because those glyphs were limited to 3–7 rays attached to a small centre circle and lacked correspondence between the represented quantity and area (Bertin 1967, p. 201). However, some experts have found it unappealing that the areas of the sectors do not increase linearly with their lengths. For instance, Bertin proposed square-root compensation for linearity. Conversely, the subjects of the psychophysical experiments were able to assess the effective dimension that coded the data without any compensation (Spence 1990, Spence and Lewandowsky 1991, Spence 2004). In general, Teghtsoonian (1965) found that subjects can selectively estimate apparent changes or physical changes depending on their understanding of the task. The former is judged exponentially ($exp = 0.8$) and the

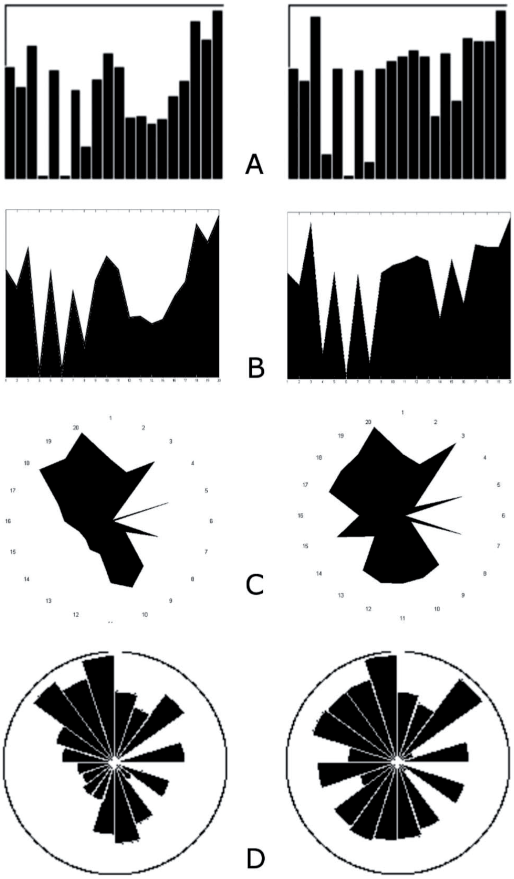

Figure 3. Four major graphs for visualising multiple dimensions of individuals: (A) bar, (B) profile, (C) star and (D) sector graph are shown. Two examples from our data are plotted for all types of graphs within each column. In graphs A and B, the dimensions are mapped from left to right and, in graphs C and D, clockwise from the top.

latter linearly. In short, sector graphs are the most viable radial alternative to bar graphs.

The hundreds of alternatives for representing multi-dimensional information (Harris 1999) are limited here by the findings of controlled experiments. We used bar and sector graphs in the experiments because they differentiate individuals and prime for independent evaluation of each dimension. Sector and profile graphs belong to studied object displays. In some real-life simulations, the judgements of nuclear plant operators and engineers were more sensitive when using star glyphs compared to bars and meters (Petersen *et al.* 1981), but the experiment used a problematic spatial design and results were statistically insignificant. Our experiments were more controlled by relying on psychophysical tradition. With the sequential comparison of different graphs in mind, special attention was paid to the timing and short-term memory requirements. Furthermore, limited exposure time can harm only local, but not global, inferences from line graphs (Carswell *et al.* 1993). The finding was explained by subjective control of attention to local inferences. We controlled this factor with very short exposure times.

5. Aims of the experiments

The following experiments evaluate the cognitive processes underlying the use of a popular service, political candidate selector (Berg *et al.* 2006). To learn from positive user feedback, the contextual components (elections) were removed to create simplified experimental settings. Thus, generalisations to other complex tasks were made possible. When 20 or 25 dimensions are reduced to locations (and clusters) on a two-dimensional plane, the dimensions lose their original meaning, which only complementary representation can reveal. These experiments focus on sector diagrams complementing reduction of dimensionality.

From the user's point of view, meaningful SOM or MDS clusters speed up the search for relevant individuals and the selection the best match. If the clustering is meaningful, adjacent individuals should be similar with respect to the original dimensions. Otherwise, the ability to focus on certain individuals is wasted. The psychophysical measures evaluate the perception of this similarity from complementary representations. Thus, the criteria for complementary diagrams are: (1) providing effective and memorable global outlines for all dimensions and (2) offering detailed values of each dimension independently. The outline needs to be memorable, because individuals need to be compared with a sequence of eye movements, not all at once (details-on-demand principle).

We operationalised the memory requirement by using a memory-based task with a two-alternative-forced-choice (2AFC) method. Our main hypothesis was that a sector graph in polar form facilitates rapid comparison of outline shapes. The speed and accuracy of subjects' performance are measured with different levels of similarity between the graph types. The sector graph is then compared to the more popular bar graph (Cartesian form), which serves the same function. Both diagrams are evaluated with presentation times above and below the time needed for saccadic eye movements. Our second hypothesis is that shorter time intervals emphasise the role of efficient outlines that ease similarity comparisons and discriminations. In the original service, related sectors were coloured with

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

similar hues. Nevertheless, very short intervals are subject to differential processing of luminance and chromatic information in the human vision system. Our third hypothesis is that black-and-white (B/W) diagrams are processed more accurately with shorter presentation times. Therefore, we needed to test all interactions between different diagram types, levels of similarity, presentation time and use of colours. Together, these parameters constitute important design features for most interactive visual interfaces.

The focus on the global outlines and temporal effects in psychophysical evaluation of graphs is a new idea. Therefore, more traditional criteria evaluate the possible trade-offs in Experiment 2. Particularly, availability of local details has been considered as a possible limitation of object displays in continuous monitoring tasks (Carswell and Wickens 1987, Wickens and Hollands 1999, p. 126, Bennet *et al.* 2000, Bennet and Walters 2001). Thus, the discrimination thresholds for relative data values on a single dimension of both diagram types are evaluated using an adaptive staircase method. For many other tasks, details are not important, so star and profile graphs are viable alternatives. Even then, the results relating to polarity are likely to be instructive.

6. Experiment 1 – holistic comparison of alternative graphs

6.1. *Subjects*

Our subjects included both of the authors and three naive volunteers (three men and two women). The authors were *blind* to the visually represented data values in both experiments and therefore had no advantage relative to volunteers. The volunteers were acquaintances who expressed interest towards the study and were compensated with two movie tickets. All subjects had several years of university studies or had graduated from university (median of age 27, range 21–52).

In general, psychophysical experiments test fewer subjects than other fields of experimental psychology, because simple visual features are better controlled and are processed with qualitatively more similar neural mechanisms. In a meta-analysis of 330 independently rated psychophysical experiments on simple visual or auditory features (published in 2007 in the *Journal of Vision, Vision Research, Perception* and *Perception and Psychophysics*) 60% had six or fewer subjects and 42% included non-naive subjects.

6.2. *Apparatus and stimuli*

Experiments were carried out in a darkened room with no light-emitting or reflecting surfaces in front of the subject except the stimulus monitor, a Dell Ultrasharp 17-inch LCD with a refresh rate of 75 Hz and 1280×1024 resolution, controlled by a Radeon 7500 graphics board. Stimuli were created with Matlab Psychophysics toolbox (Brainard 1997).

Stimuli consisted of sector and bar graphs and preparatory fixation crosses. At a viewing distance of 57 cm, one sector graph subtended a radius of 4.0 deg and one bar graph subtended 4.0 deg horizontally and 2.8 deg vertically. Both lines of the fixation cross were 40 pixels long and 1 pixel thick. When colours were used, we selected the hues for nominal categories (as suggested in Kosslyn (1989)) by choosing equidistant points in HSV-colourspace so that saturation and luminance were constant according to suggestions for non-ordered colours (e.g. Bertin 1967, p. 89). The diagrams represented the real data from the described election case, because artificial data often overemphasise the distinctiveness of the cluster structure (Siirtola 2004).

6.3. *Procedure in Experiment 1*

The procedure is illustrated in Figure 4 for sector diagram and the bar graph was treated similarly. First, a fixation cross was presented for 1.0 s. Second, it was removed and a blank screen was presented for 500 ms to reduce visual masking. The cross was made maximally thin for the same reason. Third, a standard, or reference, graph was presented for either 100 ms or 1.0 s. Fourth, immediately after the reference was removed, two alternative graphs were presented at the

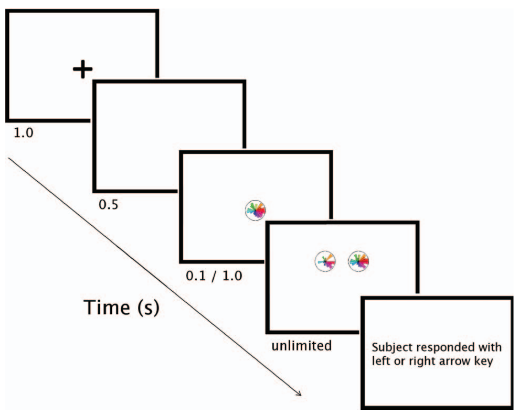

Figure 4. The five phases of the experimental procedure are illustrated for sector graphs in order of presentation (from top left to bottom right): fixation cross, blank screen, standard, alternatives, subject's response. The frames and fixation cross are thick only for illustrative purposes.

other vertices of the invisible triangle. Fifth, subjects were instructed to compare the two graphs and select the graph that was visually more similar to the standard presented earlier; subjects were to respond using the arrow key (using forefingers) on the corresponding side.

During the trials, graphs were presented in three locations that formed a downward-pointing equilateral triangle (each edge 7.0 deg): two were horizontally aligned, and the third (standard) was below and halfway between the first two. Thus, the subject could fixate on only one graph at a time with foveal vision but was able to see all graphs with the parafoveal part of the retina. A graph was either B/W or coloured and represented the data values using either length of bars or sectors. The third parameter was presentation time (100 ms or 1.0 s) of the memorised reference. The reaction time (RT) was measured from the moment the alternative graphs were presented until the subject's response. Subjects started to answer faster when they began to realise the large number of repetitions involved in the task (480 in total). Before the experiments, subjects were able to practise the task with six stimuli of each type (96 in total).

6.4. *Results of Experiment 1*

The accuracy of subjects' choices was determined by the similarity of the target graph to the presented alternatives. This is similar to some previous studies (Zacks *et al.* 1998) but a single element was used instead of the whole graph. Similarity was calculated from the city-block distance (in Minkowski metrics) of all the dimensions. Using other distance measures (SOM and Bayesian) showed similar results, but they are not reported here for simplicity. The difference between alternative graphs in their distances to the target was termed 'easiness'.

Binary categorical accuracy was submitted to the Generalised Linear Mixed Model (SAS 9.1: Glimmix procedure) with log link function and binomial error distribution. The main effects of easiness, RT, graphs (2), subjects (5), presentation times (TimeOn, 2), and possible colouring (colour, 2) were analysed (Table 1). All factors except colour had highly significant effects on the accuracy of subjects' choices. Subjects differed significantly in their performance. Easier tasks and fast RT resulted in better performance. Most importantly, all subjects performed better with sector graphs than bar graphs.

To disambiguate the main effects, all interactions with graph and individual main effects were submitted to the model (Table 2). In addition, the interaction of colour and presentation time was included. All main effects remained significant except for graphs. Of the interactions with graphs, only presentation time was significant; colour, subject, easiness and RT were not. Performance with sector graphs was better with both 100 ms and 1 s presentation times, but this effect was increased at the shorter presentation time. The B/W graphs were judged more accurately than the coloured graphs. The interaction of colour and presentation time was highly significant. At 100 ms presentation time, a relatively large difference existed in favour of B/W over coloured graphs.

Reaction times were submitted to the general linear mixed model with easiness as a continuous independent variable (Table 3). Graphs (2), subjects (5), accuracy (2), presentation times (TimeOn, 2), and possible colouring (colour, 2) were categorical independent variables. Easier choices were made faster. Subjects differed significantly in their RT. Again, the RT was faster for the correct choices. Presentation time and colour did not affect RT significantly.

For simplicity, only interactions with graphs were analysed (Table 4). There were no significant interactions with graphs and easiness. Regarding RT, the directions of effects were reported based on group means. The means themselves are not reported because they do not directly reflect the size of effects. Easiness

Table 1. Main effects for accuracy in experiment 1.

	Estimated coefficient	Standard error	Confidence interval, p
Easiness	0.07022	0.003315	<0.0001***
Subject	−0.01847	0.002117	<0.0001***
RT	−0.05407	0.005283	<0.0001***
Graph	−0.03452	0.004937	<0.0001***
Colour	−0.00040	0.004765	0.9336
TimeOn	0.04295	0.005859	<0.0001***

***$p < 0.001$.

Table 2. Main effects and interactions for accuracy in experiment 1.

	Estimated coefficient	Standard error	Confidence interval, p
Easiness	0.07492	0.00649	<0.0001***
Subject	−0.01758	0.00484	0.0003***
RT	−0.05312	0.01011	<.0001***
Graph	−0.00589	0.03888	0.8796
Colour	−0.03790	0.01401	0.0069**
TimeOn	0.02726	0.01439	0.0584
Colour × TimeOn	0.06414	0.01657	0.0001***
Graph × Colour	0.007504	0.01607	0.6406
Graph × TimeOn	−0.03345	0.01720	0.0519
Graph × Subject	−0.00088	0.00627	0.8885
Graph × Easiness	−0.00742	0.00900	0.4096
Graph × RT	−0.00039	0.01736	0.9819

**$p < 0.01$. ***$p < 0.001$.

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

Table 3. Main effects for latencies in experiment 1.

	F-value	Confidence interval, *p*
Easiness	38.387	<0.0001***
Graph	4.342	0.037*
Colour	0.684	0.408
TimeOn	2.116	0.146
Accuracy	19.034	<0.0001***
Subject	67.910	<0.0001***

*$p < 0.05$. ***$p < 0.001$.

Table 4. Main effects and interactions for latencies in experiment 1.

	F-value	Confidence interval, *p*
Easiness	39.75	$<.0001$***
Graph	3.329	0.068
Colour	0.690	0.406
TimeOn	2.253	0.133
Accuracy	19.78	<0.0001***
Subject	68.65	<0.0001***
Graph × Easiness	0.010	0.921
Graph × Colour	2.017	0.156
Graph × TimeOn	0.716	0.398
Graph × Accuracy	2.502	0.114
Graph × Subject	6.260	<0.0001***

***$p < 0.001$.

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

was not balanced between the groups but was included in the statistical models. As a result, the task was on average easier with bar graphs than with sector graphs. The difference between alternatives was the result of random sampling from real data clusters. Only the 'subjects' variable had significant interaction with graphs, resulting from longer average RTs for sector graphs by experienced subjects. When this effect was accounted for, graphs did not differ in RT. If subjects are collapsed into categories of experienced and naïve, the experienced subjects perform more accurately (coefficient $= 0.06$, $p < 0.0001$), but there is no significant interaction with graph type (coefficient $= -0.03$, $p = 0.07$) even though such as interaction exists for RT ($F = 24.5$, $p < .001$).

6.5. Discussion of Experiment 1

The main finding of Experiment 1 was that choices were more accurate with sector graphs than with bar graphs, regardless of difficulty and the subject's experience. Accuracy was reflected in the ability to select the better alternative in repeated presentations. Random sampling of real data produced easier alternatives for bar graphs in terms of measured city-block distance (1.47 vs. 1.25) and the easier alternatives were found to increase accuracy. When this is accounted for, a generalised linear model predicts better performance using sector graphs (Figure 5). Subjects varied in accuracy; however, the predicted accuracy was consistently higher (avg. 1.4%) with sector graphs for all subjects, even though the general performance level is already very high (ceiling-effect). The difference is not due to a speed–accuracy trade-off. On the contrary, accuracy increased with faster RT, because the task was easier.

The RT measure is traditional for experimental psychology and Bertin's (1967, p. 9) conception of efficiency and 'mental cost' of graphic constructs. However, the accuracy difference did not depend significantly on RT, subject, difficulty of task, or whether the graphs were coloured. Difference depended on presentation time. The accuracy benefits of using sector graphs increased with very fast display times (100 ms). At this rate, it is impossible to do any saccadic eye movements, and the diameter of the graph equals approximately the eccentricity limits of high acuity vision. Modelling simple cells early in the visual system with round receptive fields is common (e.g. Dayan and Abbott 2001). In a rapid presentation situation, perceptual integration of shape might benefit from the overall roundness of the graph. On the other hand, Kosslyn argued that 'bars often will not form perceptual groups, and it is easy to pick up individual values, but more difficult to see trends' (Kosslyn 1985, p. 510). Another difference between the two graph types is that the identity of the feature dimensions is coded by neural representations of sector graphs not only in cells tuned to retinal location (and possibly colour), but also redundantly in cells tuned to different orientations.

According to previous findings for magnitude estimation of bar lengths (Zacks *et al.* 1998), reliance on memory decreases performance accuracy and amplifies the harmful effects of the previous context. Previously, memory-based estimate of familiar areas were underestimated (with exp. $= 0.60$), whereas familiar distances were perceived rather accurately. In different experiments, unfamiliar distances were underestimated (with exp. $= 0.70$ in Radvansky *et al.* (1995)). The expected result is that unusual graphs are memorised as though more typical, but this should harm quantity estimation more than selection of a better match. Indeed, the overall accuracy in this experiment was high (Figure 5). Furthermore, decisions were more accurate with B/W graphs than with coloured graphs, especially with 100 ms presentation times, as in previous studies (Keller *et al.* 2006).

The question has been raised whether subjects in experimental visualisation research settings use only the assumed perceptual processes. The knowledge that one is being evaluated might encourage the use of higher cognitive processes (Cleveland and McGill

1984). Counter-intuitively, this has been shown to increase the variability of responses (Levine *et al.* 1996), an effect we would not observe in everyday judgements. The instructions for Experiment 1 emphasised subjective opinion. After the experiment, subjects were questioned about the experiment and their strategies. None of the naive subjects were able to describe how the quality of their performance could be evaluated. Thus, this procedure was not subject to the obvious harmful side-effects of quantity estimation research. Novice users are often unable to benefit from many visualisation techniques and critics argue that claims that novices did not have enough time to get accustomed to the use is a weak defence (van Wijk 2006). Accordingly, the validity of complicated new visualisation techniques has been questioned. This is not the case here; the difference between sector and bar graphs was consistent across different levels of experience.

7. Experiment 2 – finding and comparing one differentiating dimension

Subjects, equipment, and the presented graphs were the same as in the previous experiment.

7.1. *Procedure in Experiment 2*

Subjects were instructed to compare the aligned graphs to find the differentiating sector or bar, and to respond with an arrow key (using a forefinger) on the side that corresponded with the standard below (Figure 6). In each trial, three sector or bar graphs formed an equilateral triangle pointing down (7 deg each edge), two of them horizontally aligned and one below and halfway between them (standard). The two aligned graphs were duplicates except for one randomly selected sector or bar, for which small offset was added or reduced. If length resulting from addition of the offset exceeded 1, then reduction was performed. Also, the length was made positive.

Subjects were told that there was a 15-s time limit. When the time limit was exceeded or the subject gave an incorrect answer, a beep followed and the next trial was made easier, with larger differences between the compared graphs. If the subject gave three consecutive correct answers to trials with different (in terms of data values) graphs of the same level of difficulty, the next trial was made more challenging. This procedure followed the adaptive staircase method (Wetherill and Levitt 1965) for measuring perceptual threshold. This three-to-one rule converges to a level of 79.4% correct responses. The procedure was repeated until there were eight turning points for changing the difficulty. Step sizes of 1.5% (of maximum length) for sector graphs and 2.5% for bar graphs were selected according to pilot studies for optimising the range and sensitivity of the preselected stimuli. Sector graphs were presented first, followed by bar graphs; both were then repeated with different data values. Before the experiment, subjects were able to practise the task once for the whole procedure using both graph types.

7.2. *Results of Experiment 2*

One of the five subjects (an author) was eliminated from the analysis because he was using an exceptional strategy to fuse retinal images of the graphs. This requires much training because the graphs were 7 deg apart. Certainly, this cannot be an ecologically valid model for real use of visualisations. No other subject

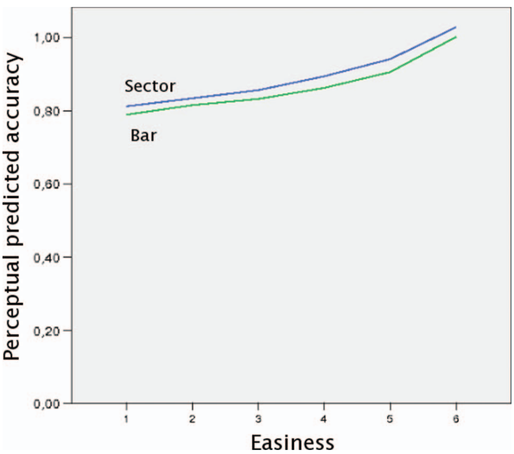

Figure 5. Percent predicted accuracy (on *y*-axis) as a function of six categories of easiness distance (on *x*-axis) shows a surprisingly consistent benefit for using sector graphs (blue line) instead of bar graphs (green line). Standard deviations were too small (0.002–0.005%) to be visible in the figure.

Figure 6. The alternative stimuli of Experiment 2 are presented here with the emphasised frames.

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

was using this strategy. They reported (during and after the experiments) that they attempted to first locate the deviant component before responding. The measured thresholds were submitted to a complete repeated measures analysis of variance (SPSS 14) with two factors (two graphs × four subjects). There was a significant main effect of graph ($F(1,56) = 84,75$, $p < 0.001$). Sector graphs had a lower threshold than bar graphs (2.9% vs. 3.4%) consistently for all four subjects, even though the subjects differed significantly ($F(3,56) = 3.82$, $p = 0.02$) and there was a significant interaction between subjects and graphs ($F(3,56) = 6.50$, $p = 0.001$).

7.3. *Discussion of Experiment 2*

The relative, not absolute, discriminability of marks (Kosslyn 1989) was measured because the difficulty of most multi-dimensional decision tasks depends on perception of relations (see Section 3.2). The previous experiments on local details have mostly involved *point reading* of absolute values (Carswell and Wickens 1987, Wickens and Hollands 1999, p. 125–129, Bennet *et al.* 2000, Bennet and Walters 2001). However, if the accuracy on absolute values is important for the task, the most appropriate solution is to include numeric representation (Bennet *et al.* 2000). Keeping that in mind, we have questioned the ecological validity of tasks measuring absolute values (Section 3.2). Furthermore, low-level perceptual processes of comparing simple separated objects seem to rely on extracted relations, instead of absolute values (Danilova and Mollon 2003). Our alternative local criterion to measure focused attention was a similarity comparison of two graphs with respect to deviant dimension. Similar to previous experiments on local tasks, our criterion implicitly required subjects to locate the deviating dimension and this was confirmed by self-reports. Experiment 2 measured just-noticeable-difference (JND; Stevens 1961) in the length of a bar or sector, not the shortest perceivable sector or bar.

This experiment confirmed that the more accurate performance with sector graphs compared to bar graphs in the holistic comparison (exp. 1) did not result from a trade-off with lower accuracy in local level differentiation. On the contrary, the use of sector graphs resulted in more accurate perceptual evaluation of represented data in both tasks. The two experiments differed with respect to required attention to visualised data dimensions (integration or discrimination) and required memory processes (standard visible or not). However, as no trade-off was observed, there was less incentive to isolate the influences of these two factors. Moreover, both tasks required memory performance of same order-of-magnitude if one utilises mostly foveal vision and saccades in Experiment 2. In fact, all naive subjects were visually observed to saccade between graphs.

The experiment was not designed to test the guidelines of linear versus exponential (area) mapping from data value to sector length. However, the JND for sector graphs was relatively independent of the baseline length of each sector, which supports the idea of linear mapping.

8. General discussion

8.1. *Validity of the evaluation criteria*

The most important finding of our experiments was that the objective similarity in data represented by sector or bar graphs predicted the subjective estimates of similarity well. Moreover, the subjects were capable of encoding the reference graph in less time than was needed to initiate an eye movement. It is therefore unlikely that they were using individual bars or sectors (dimensions) for the comparison. All together, our results are indicative of pre-attentive perception of shapes that integrate several dimensions, possibly independent of other higher-level cognitive activity (conscious or symbolic). The results thus validate three aspects of visualisations. First, it confirms (1) the global outline as an effective evaluation criterion for interactive visualisations of large quantities of data. In object displays, the global outline facilitates integration of different features. Thus, perceptual integration empowers intuition to surpass the limitations of conscious processing (cf. the customers of IKEA in a study by Dijksterhuis *et al.* (2006)).

The results also determine the criteria for (2) memorability and (3) discriminability. There was no possibility of verifying the identity of the reference graph after the alternatives were presented. The need to backtrack with gaze is very harmful because it increases the cognitive load. These short-term visual memory requirements are stressed, for instance, by the small screens of hand-held devices. The effects for cumulative performance become apparent when the rate of eye movements is compared to the rate of much less infrequent heart-beats.

Within seconds, the subjects in the second task were able to locate a small difference in any of the dimensions (discriminability); therefore, both graphs satisfied all the task requirements (1–3) for complementary interaction of decision aids. The purpose of the decision aids was to facilitate evaluation of the statistical results of multi-dimensional visualisations (with complicated algorithms) and their usefulness for decision-making. For instance, if algorithmic processing is order-preserving, the outlines of adjacent objects should appear to be similar. This purpose need not be explicit for the user, who simply explores details about individuals.

8.2. *Design implications*

The fast presentation (100 ms) simulated eye fixations and highlighted: the advantage of sector graphs relative to bar graphs, and the higher number of errors made with coloured rather than B/W graphs.

We hope that further research with direct eye movement monitoring will confirm these results.

Alternative diagrams should be considered when individual dimensions and detailed values are not important (Shah *et al.* 1999, Zacks and Tversky 1999). If the advantage of the sector graph resulted from its polar coordinate system, then a star graph should also outperform Cartesian profile graphs. Furthermore, a designer can emphasise global aspects by increasing the size of the stimulus (Navon 1977, Kinchla and Wolfe 1979, Antes and Mann 1984). However, the size of a visual chunk should not exceed 5 deg (Casey and Wickens 1986, Tullis 1997). If the accuracy of detailed values is very important, a reference frame should be used (Cleveland and McGill 1984) by drawing a box around bars or a circle around sectors to mark the maximum value. When the frame was present, absolute length was estimated more accurately for long bars than for intermediate ones (Zacks *et al.* 1998). If very small values are of importance, a dot chart is suggested (Cleveland 1984). Usually this is not the case, and normalising the data is helpful.

The order of dimensions is also important for the global outline. In particular, data-driven clustering provides a smoother silhouette for the mentioned graphs and reduces line crossings with parallel coordinates (Hurley 2004). Nonetheless, dimensions often have meaningful ad-hoc clustering.

8.3. *Conclusion*

We have established the importance of two new evaluation criteria for the interactive use of graphs: global outline and memorability. For one, the psychophysical experiments presented show that there is relevant information available in the global outline at the level of individual graphs. Second, these outlines provide a visual indicator about the validity of algorithmic reduction of dimensionality (e.g. MDS or SOM). Thus, they improve trustworthiness and intelligibility of multi-dimensional visualisations. Third, information provided by global outline aids the perceptual integration across the numerous details (here: data dimensions) and this is the most common difficulty with visualisations. In addition, our experiments show that the information from memorable outline can be utilised even within the shortest eye fixations. To conclude, the most important thing in psychophysical evaluation is to measure relevant characteristics of the real tasks.

Acknowledgement

This work was supported by grants from the Academy of Finland (No. 210676) and Helsingin kauppakorkeakoulun Tukisäätiö (No. 200821135) to I.K.

Note

1. A rare (see Petersen *et al.* 1981, Casey and Wickens 1986, Legge *et al.* 1989) but more sophisticated alternative for measuring the quality of graphs is to apply signal-detection theory.

References

Anderson, E., 1957. A semigraphical method for the analysis of complex problems. *Proceedings of the National Academy of Sciences*, 13, 923–927.

Antes, J. and Mann, S., 1984. Global-local precedence in picture processing. *Psychological Research*, 46, 247–259.

Ariely, D., 2001. Seeing sets: representation by statistical properties. *Psychological Science*, 12, 157–162.

Bauer, M. and Johnson-Laird, P., 1993. How diagrams can improve reasoning. *Psychological Science*, 4, 372–378.

Bays, P.M. and Husain, M., 2008. Dynamic shifts of limited working memory resources in human vision. *Science*, 321, 851–854.

Bennet, K.B., *et al.*, 2000. An empirical comparison of alternative methodologies for the valuation of configural displays. *Human Factors*, 42, 287–298.

Bennet, K.B. and Walters, B., 2001. Configural display design techniques considered at multiple levels of evaluation. *Human Factors*, 43, 415–434.

Beniger, J.R. and Robyn, D.L., 1978. Quantitative graphics in statistics: a brief history. *The American Statistician*, 32, 1–11.

Berg, M., *et al.*, 2006. Exploring political agendas with advanced visualization and interface tools. *e-Service Journal*, 4, 47–63.

Berry, D.C. and Broadbent, D.E., 1988. Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology*, 79, 251–272.

Bertin, J., 1967. *Semiology of graphs*. Madison, WI: University of Wisconsin Press.

Bertin, J., 1981. *Graphics and graphic information-processing*. Berlin: de Gruyter.

Biederman, I., 1972. Perceiving real-world scenes. *Science*, 177, 77–80.

Biederman, I., 1982. Scene perception: detecting and judging objects undergoing relational violations. *Cognitive Psychology*, 14, 143–177.

Biederman, I., Glass, A., and Stacy, W., 1973. Searching for objects in real-world scenes. *Journal of Experimental Psychology*, 97, 22–27.

Biederman, I., Rabinowitz, J., and Stacy, W., 1974. On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, 103, 597–600.

Brainard, D.H., 1997. The psychophysics toolbox. *Spatial Vision*, 10, 433–436.

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

Card, S., Mackinlay, J., and Schneiderman, B., eds. 1999. *Readings in information visualization: using vision to think*. San Francisco, CA: Morgan Kaufmann Publishers.

Carlson, T. and Strandberg, K., 2005. The European Parliament Election on the Web: Finnish Candidates on-line: supply and demand. *In*: *17th Nordic Conference on Media and Communication Research*. Aalborg, Denmark: Society of Media Researchers.

Carpenter, P.A. and Shah, P., 1998. A model of perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, 4, 75–100.

Carswell, C.M., Emery, C., and Lonon, A.M., 1993. Stimulus complexity and information integration in the spontaneous interpretations of line graphs. *Applied Cognitive Psychology*, 7, 341–357.

Carswell, C.M. and Wickens, C.D., 1987. Information integration and the object display: an integration of task demands and display superiority. *Ergonomics*, 30, 511–527.

Casey, E.J. and Wickens, C.D., 1986. *Visual display representation of multidimensional systems: the effect of information correlation and display integrality*. Technical Report, University Of Illinois, CPL-86-2/MDA903-83-K-0255.

Chambers, D. and Reisberg, D., 1985. Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception and Performance*, 11, 317–328.

Chambers, J., *et al*., 1983. *Graphical methods for data analysis*. Pacific Groove, CA: Wadsworth & Brooks.

Chernoff, H., 1973. The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*, 69, 361–368.

Cleveland, W., 1984. Graphical methods for data presentation: full scale breaks, dot charts, and multibased logging. *The American Statistician*, 38, 270–280.

Cleveland, W. and McGill, R., 1984. Graphical perception: theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79, 531–554.

Cleveland, W. and McGill, R., 1985. Graphical perception and graphical methods for analyzing scientific data. *Science*, 229, 828–833.

Cowan, N., 2000. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–185.

Cowan, N. and Morey, C., 2006. Visual working memory depends on attentional filtering. *Trends in Cognitive Sciences*, 10, 139–141.

Cox, R., 1999. Representation construction, externalised cognition and individual differences. *Learning and Instruction*, 9, 343–363.

Cozby, B., 2004. *Methods in behavioral research*. New York: McGraw-Hill Co.

Danilova, M. and Mollon, J., 2003. Comparison at a distance. *Perception*, 32, 395–414.

Dayan, P. and Abbott, L.F., 2001. *Theoretical neuroscience: computational and mathematical modeling of neural systems*. Cambridge, MA: MIT Press.

Dijksterhuis, A., *et al*., 2006. On making the right choice: the deliberation-without-attention effect. *Science*, 311, 1005–1007.

Duncan, J., 1984. Selective attention and organization of visual information. *Journal of Experimental Psychology: General*, 113, 501–517.

Eells, W.C., 1926. The relative merits of circles and bars for representing component parts. *Journal of American Statistical Association*, 21, 119–132.

Gärdenfors, P., 2000. *Conceptual spaces: the geometry of thought*. Boston, MA: MIT Press.

Gattis, M. and Holyoak, K., 1996. Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 231–239.

Gilhooly, K.J., *et al*., 1993. Working memory and strategies in syllogistic-reasoning tasks. *Memory and Cognition*, 21, 115–124.

Guthrie, J.T., 1988. Locating information in documents: examination of a cognitive model. *Reading Research Quarterly*, 23, 178–199.

Guthrie, J.T., Weber, S., and Kimmerly, N., 1993. Searching documents: cognitive processes and deficits in understanding graphs, tables, and illustrations. *Contemporary Educational Psychology*, 18, 186–221.

Hahn, G.J., Morgan, C.B., and Lorensen, W.E., 1983. Color face plots for displaying product performance. *IEEE Computer Graphics and Applications*, 3, 23–29.

Harris, R., 1999. *Information graphics: a comprehensive illustrated reference*. Oxford: Oxford University Press.

Hurley, C., 2004. Clustering visualizations of multidimensional data. *Journal of Computational and Graphical Statistics*, 13, 788–806.

Hutchins, E., 1995a. *Cognition in the wild*. Cambridge, MA: MIT Press.

Hutchins, E., 1995b. How a cockpit rembembers its speed. *Cognitive Science*, 19, 265–288.

Inselberg, A., 1999. Multidimensional detective. *In*: S.K. Card, J.D. Mackinlay, and U.S. Schneiderman, eds. US: Academic Press, 107–114.

Inselberg, A. and Dimsdale, B., 1990. Parallel coordinates: a tool for visualizing multi-dimensional geometry. *Proceedings of the 1st conference on Visualization, '90*, 361–378.

Jacoby, W., 1998. *Statistical graphics for visualizing multivariate data*. London, UK: Sage Publication Inc.

Kaipainen, M., *et al*., 2001. Presenting data as similarity clusters instead of list. Data from local politics as an example. *In*: C. Stephanidis, ed. *Universal access in HCI towards an information society for all*. London: Lawrence Erlbaum.

Keller, T., *et al*., 2006. Information visualization for knowledge acquisition: the impact of dimensionality and color coding. *Computers in Human Behavior*, 22, 43–65.

Kinchla, R. and Wolfe, J., 1979. The order of visual processing: "Top-down", "bottom-up", or "middle-out". *Perception and Psychophysics*, 25, 225–231.

Kleinmuntz, B., 1990. Why we still use our heads instead of formulas: toward an integrative approach. *Psychological Bulletin*, 107, 296–310.

Kleinmuntz, D.N. and Schkade, D.A., 1993. Information displays and decision process. *Psychological Science*, 4, 221–227.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59–69.

Kohonen, T., 2001. *Self-organizing maps*. 3rd ed. NY: Springer-Verlag.

Kontkanen, P., *et al*., 2000. Supervised model-based visualization of high-dimensional data. *Intelligent Data Analysis*, 4, 213–227.

Kosslyn, S., 1985. Review: graphics and human information processing: a review of five books. *Journal of the American Statistical Association*, 80, 499–512.

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012

Kosslyn, S., 1989. Understanding charts and graphs. *Applied Cognitive Psychology*, 3, 185–225.

Kramer, A.F., Wickens, C.D., and Donchin, E., 1985. Processing of stimulus properties: evidence for dual-task integrality. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 393–408.

Laine, S., 2003. *Using visualization, variable selection and feature extraction to learn from industrial data*. Thesis (PhD). Helsinki University of Technology.

Lappin, J.S., 1967. Attention in the identification of stimuli in complex visual displays. *Journal of Experimental Psychology*, 75, 321–328.

Larkin, J. and Simon, H., 1987. Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65–99.

Legge, G., Gu, Y., and Luebker, A., 1989. Efficiency of graphical perception. *Perception and Psychophysics*, 46, 365–374.

Leinhardt, G., Zaslavsky, O., and Stein, M., 1990. Functions, graphs, and graphing: tasks, learning and teaching. *Review of Educational Research*, 60, 1–64.

Levine, G., Halberstadt, J., and Goldstone, R., 1996. Reasoning and the weighting of attributes in attitude judgments. *Journal of Personality and Social Psychology*, 70, 230–240.

Luck, S. and Vogel, E., 1997. The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.

Moring, T. and Mykkänen, J., 2005. Voter choice and political mediation: net-based candidate selectors in Finnish elections. *In*: *The 55th annual conference of the International Communication Association*. New York: International Communication Association.

Navon, D., 1977. Forest before trees: the precedence of global features in visual perception. *Cognitive Psychology*, 9, 353–383.

Neisser, U., 1976. *Cognition and reality*. San Francisco: WH Freeman.

Petersen, R.J., Banks, W.B., and Gertman, D.I., 1981. Performance-based evaluation of graphic displays for nuclear power plant control rooms. *In*: *Conference on Human Factors in Computer Systems*. New York: ACM Press, 182–189.

Potter, M., 1976. Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 509–522.

Potter, M., 1993. Very short-term conceptual memory. *Memory and Cognition*, 21, 156–161.

Potter, M., *et al.*, 2002. Recognition memory for briefly presented pictures: the time course of rapid forgetting. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1163–1175.

Pomerantz, J.R., 1981. Perceptual organization in information processing. *In*: M. Kubovy and J.R. Pomerantz, eds. *Perceptual organization*. Hillsdale, NJ: Lawrence Erlbaum, 141–180.

Poulton, E.C., 1985. Geometric illusions in reading graphs. *Perception and Psychophysics*, 37, 543–548.

Radvansky, G.A., Carlson-Radvansky, L.A., and Irwin, D.E., 1995. Uncertainty in estimating distances from memory. *Memory and Cognition*, 23, 595–606.

Reber, A.S., 1976. Implicit learning of synthetic languages: the role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 88–94.

Robertson, G., Card, S., and Mackinlay, J., 1989. The cognitive coprocessor architechture for interactive user interfaces. *In*: *2nd annual ACM SIGGRAPH symposium of User Interface software and technology*. New York: ACM Press, 10–18.

Scaife, M. and Rogers, Y., 1996. External cognition: how do graphical representations work? *International Journal of Human–Computer Studies*, 45, 185–213.

Schooler, J.W. and Engstler-Schooler, T.Y., 1990. Verbal overshadowing of visual memories: some things are better left unsaid. *Cognitive Psychology*, 22, 36–71.

Schwartz, S.H., 1971. Modes of representation and problem solving: well evolved is half solved. *Journal of Experimental Psychology*, 91, 347–350.

Shah, P., Mayer, R.E., and Hegarty, M., 1999. Graphs as aids to knowledge construction: signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology*, 91, 690–702.

Shanks, D., 2006. Complex choices better made unconsciously? *Science*, 313, 760–761.

Shepard, R., 1962. Analysis of proximities: multidimensional scaling with an unknown distance function I & II. *Psychometrika*, 27, 125–139, 219–246.

Shneiderman, B., 1996. The eyes have it: a task by data type taxonomy for information visualization. *Proceedings IEEE Visual Languages*. Washington, DC: IEEE Computer Society, 336–343.

Shneiderman, B., 2002. Inventing discovery tools: combining information with data mining. *Information Visualization*, 1, 5–12.

Siegrist, M., 1996. The use or misuse of three-dimensional graphs to represent lower-dimensional data. *Behaviour and Information Technology*, 15, 96–100.

Siirtola, H., 2004. Interactive cluster analysis. *Eight International Conference on Information Visualization IV*. 4, 471–476. Washington, DC: IEEE Computer Society.

Simkin, D. and Hastie, R., 1987. An information-processing analysis of graph perception. *Journal of the American Statistical Association*, 82, 454–465.

Spence, I., 1990. Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 683–692.

Spence, I., 2004. The apparent and effective dimensionality of representations of objects. *Human Factors*, 46, 738–747.

Spence, I. and Lewandowsky, S., 1991. Displaying proportions and precentages. *Applied Cognitive Psychology*, 5, 61–77.

Spence, R., 2002. Rapid, serial and visual: a presentation technique with potential. *Information Visualization*, 1, 13–19.

Stenning, K. and Oberlander, J., 1995. A cognitive theory of graphical and linguistic reasoning: logic and implementation. *Cognitive Science*, 19, 97–140.

Stevens, S.S., 1961. The psychophysics of sensory function. *In*: A. Rosenblith, ed. *Sensory communication*. Cambridge, MA: MIT Press, 1–33.

Sweller, J., *et al.*, 1990. Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, 119, 176–192.

Teghtsoonian, M., 1965. The judgment of size. *American Journal of Psychology*, 78, 392–402.

Treisman, A., 1982. Perceptual grouping and attention in visual search for features and objects. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 194–214.

Tufte, E.R., 1983. *The visual display of quantitative information*. Cheshire, CO: Graphichs Press.
Tukey, J., 1977. *Exploratory data analysis*. Reading, MA: Addison-Wesley.
Tullis, T., 1997. Screen design. *In*: M. Helander, T.K. Landauer, and P. Prabhu, eds. *Handbook of human-computer interaction*. Amsterdam: North Holland, 503–531.
van Wijk, J., 2006. Views on visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12, 421–433.
Vecera, S. and Farah, M., 1994. Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, 123, 146–160.
Vernon, M.D., 1950. The visual presentation of factual data. *British Journal of Educational Psychology*, 20, 174–185.
Wainer, H. and Thissen, D., 1988. Plotting in the modern world: statistics packages and good graphics. *Chance*, 1, 10–20.
Ward, M.O., 2002. A taxonomy of glyph placement strategies for multidimensional data visualization. *Information Visualization*, 1, 194–210.
Wertheimer, M., 1938. Laws of organization in perceptual forms. *In*: W. Ellis, ed. *A source book of Gestalt psychology*. London: Routledge & Kegan Paul, 71–88.
Wetherill, G. and Levitt, H., 1965. Sequential estimation of points on a psychometric function. *The British Journal of Mathematical and Statistical Psychology*, 18, 1–10.
Wickens, C. and Hollands, J., 1999. *Engineering psychology and human performance*. 3rd ed. New Jersey: Prentice-Hall Inc.
Zacks, J., *et al*., 1998. Reading bar graphs: effects of extraneous depth cues and graphical context. *Journal of Experimental Psychology: Applied*, 4, 119–138.
Zacks, J. and Tversky, B., 1999. Bars and lines: a study of graphic communication. *Memory and Cognition*, 27, 1073–1079.
Zhang, J., 1997. The nature of external representations in problem solving. *Cognitive Science*, 21, 197–217.
Zhang, J. and Norman, D.A., 1995. A representational analysis of numeration systems. *Cognition*, 57, 271–295.

Downloaded by [University of Helsinki], [Mikko Berg] at 09:36 28 February 2012