

# Gravi++: Interactive Information Visualization of Highly Structured Temporal Data

Klaus Hinum<sup>1</sup>, Silvia Miksch<sup>1</sup>, Wolfgang Aigner<sup>1</sup>, Susanne Ohmann<sup>2</sup>, Christian Popow<sup>2</sup>, Margit Pohl<sup>3</sup>, and Markus Rester<sup>3</sup>

<sup>1</sup>Institute of Software Technology & Interactive Systems, Vienna University of Technology

<sup>2</sup>Department of Child and Adolescent Neuropsychiatry, Medical University of Vienna

<sup>3</sup>Design and Assessment of Technologies Institute, Vienna University of Technology

Vienna, Austria, <http://ieg.ifs.tuwien.ac.at/projects/in2vis/>

## Abstract

Tracking and comparing psychotherapeutic data derived from questionnaires involves a large number of highly structured, time-oriented parameters. Descriptive and other statistical methods are only suited for partial analysis. Therefore, we invented a spring-based interactive Information Visualization method for analysing these data more in-depth. With our method the user is able to find new predictors for a positive or negative course of the therapy due to the combination of various visualization and interaction methods.

## 1 Introduction

Visualization tools have been used in the medical domain for a long time. The majority of methods and tools was developed for the field of scientific visualization, for example 3D volume visualization tasks or computer tomography visualizations. In the field of medical treatment planning different kinds of data need to be explored in the reasoning process, such as time-oriented patient data or the course of the patient state. Information Visualization (InfoVis) techniques can be used to support this exploration process and promote a deeper level of understanding of the data and information under investigation. To complement them, task-specific interaction methods have to be developed.

We are aiming for supporting a psychotherapeutic study on anorexic girls where high dimensional, abstract, time-oriented medical data is collected. The analysis of these data is a challenging process. One way to deal with this problem would be to apply statistical methods. However, these methods are limited to prove known hypotheses and can hardly capture the complex process of therapeutic interventions with not yet discovered hypotheses. To overcome this limitation, we are investigating interactive information exploration techniques.

We have developed a new interactive InfoVis method, called Gravi++<sup>1</sup>, to observe new interdependencies between various kinds of parameters. Gravi++ uses the capabilities of the human perceptual system by displaying moving icons on the screen following a spring-based model to

<sup>1</sup>The name Gravi++ is a mixture of two metaphors. “Gravi” stands for gravitation and the two summation signs stand for two magnetic plus poles.

facilitate the reasoning process. Furthermore, this process is strongly supported by task-specific interaction methods.

## 2 The Medical Problem

Gravi++ is intended to analyze questionnaires of girls with eating disorders (Anorexia Nervosa (AN)). At the Department of Child and Adolescent Neuropsychiatry, Medical University of Vienna, a study is taking place, in which alternative therapeutic processes (cognitive behavioral therapy) of anorexic girls are compared. The psychologists and physicians who are working with the girls need to explore the data in an experimental way in order to derive time-oriented quantitative and qualitative information on the states of the patients. The important features of the involved data structures are their data types, complexity (highly structured), and temporal dimensions.

### Data Characteristics

Every patient, their parents, and their therapists have to answer an extensive set of questionnaires before, during (three-monthly), and after the therapy (each time 20 questionnaires per patient, four per parent and one per therapist). Each questionnaire consists of about 40 questions. An additional challenge is that the questionnaires are not all asked in the same interval. Some are even used only once. After a normalization process the data derived from these questionnaires ranges from '0' to '6'. This range can mean 'yes' to 'no', but can also stand for nominal values like 'feeling guilty after eating a meal'.

Example questions of different questionnaires are:  
FAMOS14: 'To relax is': (1) totally unimportant (2) unimportant, (3) does not matter, (4) rather important to me, (5) extremely important to me.

MRFSF1: 'I treat myself to tranquility and recreation': (1) no, (2) rather no, (3) rather yes, (4) yes.

EAT13: 'I feel sick after eating': (0) never, (1) seldom, (2) sometimes, (3) often, (4) very often, (5) always.

Some questions are concatenated to so called predictors. These predictors should give an idea about a specific status of a patient and is used to predict the further development of the therapy.

The following predictors have been identified during our analysis: (1) Predictors for a negative therapeutic course are 'lacking close friendships', 'comorbid personality disorder', 'low self directedness', 'lacking sincerity in answering questionnaires because of highly social desirable an-

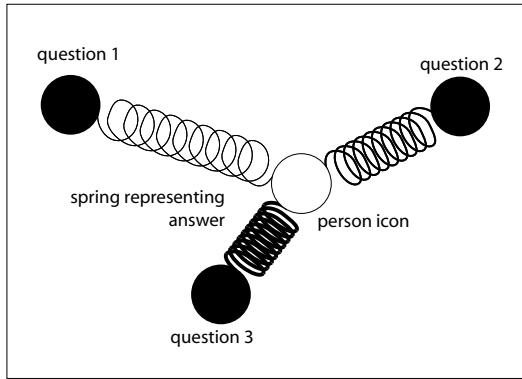


Figure 1: General Principle: A person is attracted by three questions. The answer is modelled with an invisible spring. The higher the question is answered, the stronger the spring pulls to the black disk. We plotted the springs to illustrate this general principle.

swering style’, and ‘denying disease’; (2) Predictors for positive course are a ‘decreasing depression score after half a year of cognitive behavioral therapy (CBT)’ or an ‘increasing value of feelings of pleasure by doing favoring activities after 3 months of CBT’.

The task of finding new predictors, like the ones mentioned above, lead to the development of the new visualization technique Gravi++.

### 3 Related Work

Different kinds of InfoVis methods were developed in the last years. Important features of InfoVis methods are to support the exploration process of complex, heterogeneous data promoting a deeper level of understanding of the data and information and to foster new insights into the underlying exploration process and the data themselves [Ware, 2000; K. Card, 1999].

Medical highly structured data, such as psychotherapeutic data, impose an additional challenge due to their complexity and their temporal dimensions. Worm Plots [Matthews and Roze, 1997], the Zoom Star solution [Noirhomme-Fraiture, 2002], the TimeWheel [Tominski *et al.*, 2003], the Table Lens [Rao and Card, 1994], Stardimates [Lanzenberger *et al.*, 2003b], and LinkVis [Lanzenberger *et al.*, 2003a] are techniques that try to visualize and explore such kinds of data. However, none of these techniques provide enough interaction possibilities to find new interdependencies (such as predictors) and explore the data thoroughly. Furthermore, we are dealing with a huge amount of highly structured time-oriented data, which needs appropriate methods to analyze and to discover patterns. Therefore, we created Gravi++ with the help of the following concepts.

We were inspired by the outstanding idea of the Vibe System [Hendley *et al.*, 1995] to position document icons according to the occurrence of keywords with a spring-based system and adapted that idea for our core visualization. A similar approach was proposed in the RAD-VIZ method [Brunsdon *et al.*, 1998] to map a set of m-

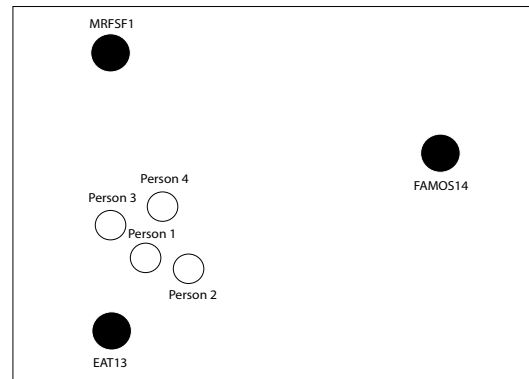


Figure 2: Simple example of finding a cluster: All persons gave similar answers to the three questions. Therefore, a cluster of all person’s icons can be seen.

dimensional points onto two dimensional space. The idea for the different metaphors was inspired by the work of [McGinn and Picking, 2003]. Animation is one way to effectively visualize temporal changes, which is shown in [Nakakoji *et al.*, 2001]. Furthermore, to present another view of multidimensional data, we used Star Glyphs such as those presented in the XmdvTool [Ward, 1994].

## 4 Gravi++

### 4.1 Concept

The human perceptual system has a remarkable ability to organize and locate things spatially, judge comparative sizes, distinguish between a large range of colors and patterns and perceive motion [Olsen *et al.*, 1993]. Gravi++ tries to utilize these human capabilities by positioning icons on the screen. There are two kinds of icons representing (1) patients and (2) questions from the questionnaires respectively (compare Figure 3). Every patient is attracted by the questions according to the answer she gave. This is modeled with a spring-based system. Every question is connected with every person by a spring. This is illustrated in Figure 1 showing a person who is attracted by three questions.

The strength of the individual springs depends on the answer the patient gave. This way, every person’s icon position on the screen identifies how she answered the involved questions. This leads to the formation of clusters of persons who gave similar answers. Because of the fact that things that are close together are perceived as a group, the finding and differentiation of clusters is an easy task for the human perceptual system according to the Gestalt Laws [Koffka, 1935]. In Figure 2 all persons gave similar answers to the questions MFRSF1, FAMOS14, and EAT13. Therefore, their icons form a cluster near EAT13. Furthermore, this tells us, that all persons have answered EAT13 with a higher value than the other questions.

The size of a person’s icon can be mapped to any additional parameter (for example to the body mass index of the patients) or to the attraction force. In the second case the sphere is larger if it is attracted by higher values. This fea-

ture helps discriminating different icons that are attracted by the same values with a different coefficient. For example, if a person has answered the questions one, two, and three with answer number one, the icon is on the same place as a person that has answered each question with answer number five.

To visualize the changing values over time, Gravi++ uses animation. The position of each person's icon change over time allowing to trace, compare and analyze the changing values. Alternatively the change over time can also be represented by traces. The size and path of the person's icon is shown corresponding to all time steps or only to a restricted subset like the previous and the next time step.

To visualize the exact values of each question, rings around the question's icon can be drawn. The ring size corresponds with the attraction to the question. To avoid overlapping rings with the same value, they are put closely side by side.

In addition, Star Glyphs [Ward, 1994] can be shown, which communicate the exact values. The edges of the Star Glyph are connected with the corresponding question rings and both are drawn in the same color as the person icon. This helps the user to identify the corresponding person.

Missing data is handled by the system in two ways. If a person has answered no questions at a specific time step, the icon of the person is transparent. If a value of the size of a person's icon is missing, the icon is shown with a special marking.

Gravi++ is intended for a restricted parameter space. The more questions are selected, the smaller is the influence of a single question on the position of the person's icons. Furthermore, too many person icons lead to clutter because of overlapping icons. To select a suitable subset of parameters, we have implemented a list-based overview visualization.

The main aim of the visualization is to derive predictors. In the following, different kinds of interactions are explained to support that task.

## 4.2 Interactions with Gravi++

Gravi++'s main intention is to provide functionalities to get new insights into the data, like finding clusters or similarities in the movement over time. Furthermore, icons that are drifting apart can give important clues regarding the data. For this purpose Gravi++ provides a set of interactions.

All interactions can be classified in three categories: (1) interactions on question's icons, (2) interactions on person's icons, and (3) general interactions. In the following these categories are explained in detail.

### Interactions on Question's Icons

- *Add or remove a question.*
- *Change the position of a question:* The position can be changed either freely or arranged on a circle. Furthermore, the user can choose to evenly space the distance between the questions on the circle. When changing the position of the questions, the corresponding positions of the person's icons change simultaneously. With this feature the user can interactively change the positions of the questions to find new clusters or other interesting visual structures. The traces and the

Star Glyph also change their form automatically when dragging a question around. This feature is called 'live preview'.

- *Change the influence of a question* (see  $strength_j$  in subsection 4.3): This enables the user to emphasize the influence of a chosen question.
- *Add complementary question:* By executing this option on a question, a new question with complementary values is created.
- *Highlighting a question:* By highlighting a question, only the highlighted question's rings are shown. If no question is highlighted, all rings are shown.
- *Hide question icons:* All question icons can be hidden to reduce clutter and facilitate the analysis of traces or the Star Glyph.

### Interactions on Person's Icons

- *Add or remove a person's icon.*
- *Change the parameter representing the sphere size:* You can map the parameters of a single question, the power of attraction, or an increasing size over time to the sphere size.
- *Change the speed of movement:* Adjust how fast the person icons are moving to their destination.
- *Hide person icons:* All person icons can be hidden to reduce clutter and facilitate the analysis of traces or Star Glyphs.
- *Show traces:* The user has the option to show traces ranging over all time steps or only a restricted subset.

### General Interactions

- *Save and load current settings and visualization:* The position of all elements and all settings can be saved for later analysis or logging purpose.
- *Show Star Glyph:* The user can superimpose a Star Glyph of the currently displayed data set to clarify the actual values. The corners of the Star Glyph correspond to the positions of the questions.
- *Inverse Star Glyph:* The edges of the Star Glyph are painted either to the center of the visualization or outwards.
- *Next and previous time step function:* This changes the time parameter to the next or previous one.
- *Direct selection of time steps:* The user can directly choose the time step on a time line or in a list.
- *Highlighting:* Persons and questions can be highlighted.
- *Show grid:* The user may fade in a grid in the background. This helps judging distances and sizes.

## 4.3 Algorithm

As we explained the idea and concept of Gravi++ in section 4.1, now we present the algorithm for positioning the person's icons. The force from each person  $i$  and each question  $j$  for each axis (in our case for the x and y axis) is  $xforce_i = \sum_{j=0}^n strength_j \cdot value_{ij} \cdot f(qx)$

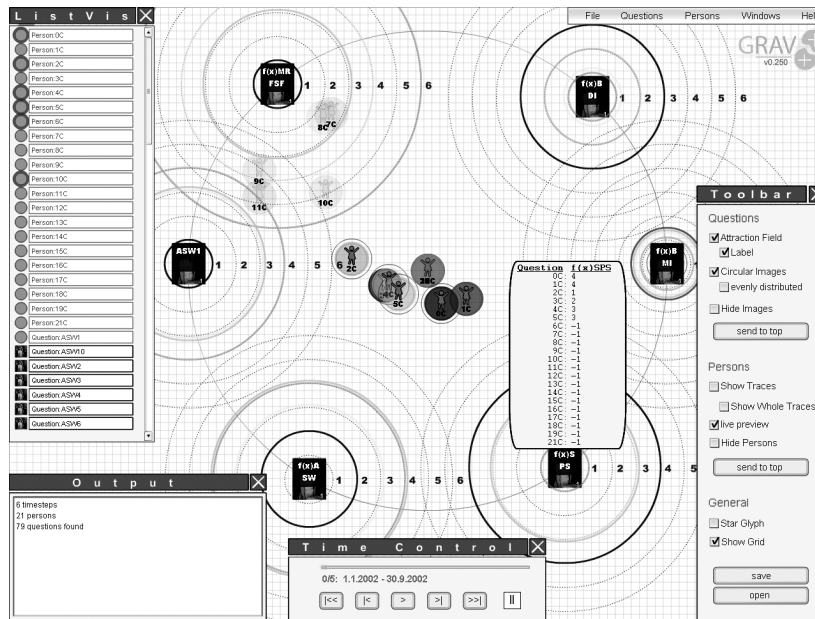


Figure 3: Typical application case with Gravi++.

and  $yforce_i = \sum_{j=0}^n strength_j \cdot value_{ij} \cdot f(qy)$ , where  $strength_j$  stands for the manually set strength multiplier of the question  $j$  (can be altered in the user interface and ranges from 0 = no attraction to  $n$  = an arbitrary value).  $value_{ij}$  stands for the answer of person  $i$  to question  $j$ .  $f(qx)$  and  $f(qy)$  stands for the attraction over the distance for each axis. Currently, we use a linear function that grows stronger over the distance:  $f(qx) = qx_j - px_i$ ,  $f(qy) = qy_j - py_i$ . Here  $qx_j$  and  $qy_j$  stand for the x, y coordinates of the question  $j$  and  $px_i$  and  $py_i$  stand for the x, y coordinates of the person  $i$ .

Now we can analytically solve the equations by substituting the force with zero ( $xforce_i = 0$  and  $yforce_i = 0$ ), to find the final position for each person ( $px_i, py_i$ ).

#### 4.4 Implementation

Gravi++ was implemented as a prototype during the in2vis project<sup>2</sup>. The prototype was implemented in Macromedia Flash MX 2004 because of its rapid visual development possibility.

The system consists of two visualizations that work closely together. There is an overview visualization (ListVis), to select a subset of a large data set. This subset can then be analyzed with the main visualization Gravi++. The data exchange between these two modules is implemented by drag and drop. You can simply drag a person or question to Gravi++ and explore it further. Both modules support linking and brushing.

Gravi++ was implemented with different metaphors which can be exchanged. In the first of the currently two implemented metaphors the persons are symbolized by iron spheres that are attracted by magnets standing for the questions. The circles around the magnets stand for the mag-

netic fields and visualize the concrete answers. The second metaphor shows the persons as people and maps the questions to paintings in an art exhibition. Here the metaphor is explained by persons who are attracted by beautiful paintings. The person moves closer to those paintings she likes. The circles around the images stand for how much the person likes the image or not. Other metaphors can be implemented by exchanging the icons representing persons and questions.

The persons are color-coded with twelve distinct colors as Colin Ware proposed in [Ware, 2000]. To enable the user to recognize icons that lie on top of each other, all icons are drawn slightly transparent. To control the time steps, a time line was implemented showing the current time step and providing the “video-like” controls *play*, *rewind*, *stop*, *go to first/last frame*, *go to next/previous frame*. Tool tips give complete information on the data item pointed at with the mouse.

In Figure 3 you can see a screenshot of a typical application case.

### 5 Example: How to Find a Predictor

Predictors allow to assess the development of parameters after a certain treatment step had been performed. Therefore, we have to analyse the differences in the first answers and compare them with the differences over all timesteps. This can be done by watching the animations over time or analysing the plotted traces of the persons’s icons. Furthermore, the change over all timesteps can also show some sort of pattern that indicates a new predictor.

Experimenting sessions with our partners of the Department of Child and Adolescent Neuropsychiatry identified the following steps to find predictors. We will illustrate these steps using an example of our dataset consisting of nine patients and 73 questions in six time steps:

<sup>2</sup>Interactive Information Visualization: Exploring and Supporting Human Reasoning Processes

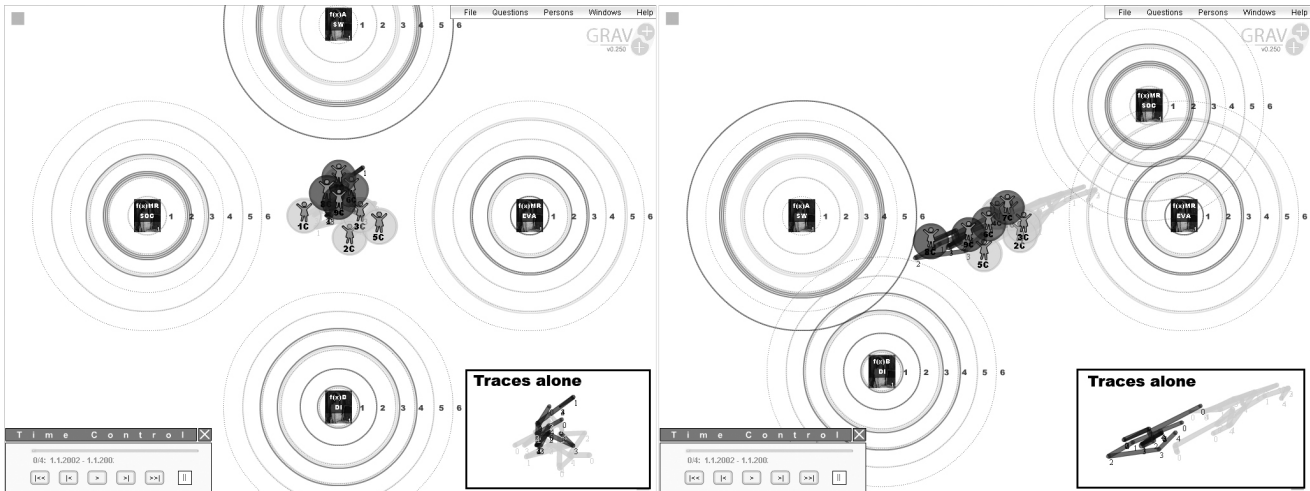


Figure 4: The left-hand screenshot shows four questions and no discernable pattern in the traces. The right-hand screenshot shows the same questions ordered differently with two easily distinguishable clusters in the traces.

1. We choose the questionnaires ASW (inverted)<sup>3</sup>, BDI<sup>4</sup>, MR EVA<sup>5</sup>, MR SOC<sup>6</sup> and all persons from the overview visualization and add them to Gravi++.
2. With the enabled option “Full Traces” we can analyse the traced paths of the persons over time.
3. By moving around the questions we try to find clusters of traces. In our example we have nine different persons. They are color coded the following way: The icons and traces with dark grey shades had not a favourable outcome in the therapy. The icons and traces with light grey shades had a positive therapy progress. This means we have to find clusters of evenly coloured traces.
4. By positioning MR EVA and MR SOC on the one side and ASW (inverted) and BDI on the other side we can build two clusters of light and dark traces. In the left screenshot on Figure 4 you can see the four questions before positioning (with no identifiable clustering of the traces) and on the right screenshot afterwards (with two identifiable clusters of light and dark traces).
5. By moving through the timesteps we can see that the randomly located persons in the first timestep, compose two clusters in the second and all later timesteps. Therefore, we can conclude that high values of BDI and the inverted ASW after three months therapy predict a bad outcome of the therapy, whereas high values

<sup>3</sup>ASW: The Self-Efficacy Scale Index is a summary of the test items assessing self-esteem and strategies to cope with difficult situations independently.

<sup>4</sup>BDI: The Beck-Depression-Index describes the severity of depression.

<sup>5</sup>MR EVA: The Marburg Inventory Index specifies the feeling of pleasure in hedonistic behavior.

<sup>6</sup>MR SOC: The Sense of Coherence Scale Index indicates the intensity of comprehensibility, manageability and meaningfulness as the internal psychological mechanism mediating the effects of external stressors and resources on psychological dysfunction.

of MR EVA and MR SOC predict a good outcome. Furthermore, we verify this claim by watching the movement over all remaining timesteps. In our example the persons with a positive therapy outcome still move further to the pole with MR EVA and the MR SOC and the persons with a negative outcome to the other pole. This means we can conclude that low BDI values and high MR EVA, MR SOC, and ASW values could be a predictor for a positive therapy outcome. This is confirmed by the clinical impression that patients who respond more rapidly to the primary therapeutic goals (enjoying pleasure and being in a good mood), have a better outcome.

## 6 Benefits and Limitations

The visual elements of Gravi++ use several advantages from a cognitive perspective [K. Card, 1999]. Especially, the interactive manipulation can help the user get new insights through the data. This can be used to formulate and test a hypothesis on the data e.g., to find a new predictor. The combination of different visualization techniques, like Star Glyph, traces, an overview visualization, and the Gravi++ core itself increases the possibilities to find new insights. What is special about Gravi++ is the combination of these advantages and its orientation on medical data.

Nevertheless, some problems and shortcomings of Gravi++ are still not solved. Incomplete data leads to incomparable person icon positions because there is no attraction from questions that were not answered. A solution to this problem could be to use the value from the last time step, to use an average value, or a default value. A restriction of Gravi++ is the parameter space. Too many questions lead to clutter and make the interactive change very difficult. The impact of one question on the person’s position declines the more questions are shown. Furthermore, if too many persons are analyzed, many icons would overlap and it would be difficult to interact with the representation. The values of the rings surrounding each question would be in-

distinguishable. This happens because rings representing the same value are not shown on top of each other but next to each other. The solution for such situations is to use highlighting of subgroups for detailed exploration.

## 7 Conclusion and Future Work

We have presented an interactive InfoVis method, called Gravi++, which addresses the particular features of abstract, highly structured data which are acquired during cognitive behavioral treatment (CBT) of anorexia nervosa in adolescent girls. This data is difficult to explore by descriptive and other statistical methods. Because of Gravi++'s various visualization and interaction techniques it is an appropriate method for finding new predictors in the data. Our cooperating psychologists see various application areas for this InfoVis technique within their clinical study of anorexic girls.

A possible extension to Gravi++ could be an algorithm that automatically positions the questions to find clusters of persons.

In the next step we plan to integrate the results of an in-depth user interface study with about 20 participants. Afterwards, a large number of subjects will be involved in a study, that will compare Gravi++ with supervised machine learning and exploratory data analysis to get new insights on the impact on the human reasoning process.

## Acknowledgements

This project is supported by The Vienna Science and Technology Fund ('Wiener Wissenschafts-, Forschungs- und Technologiefonds' - WWTF), grant WWTF CI038. Many thanks to the anonymous reviewers who helped us to improve our paper.

## References

- [Brunsdon *et al.*, 1998] C. Brunsdon, A. S. Fotheringham, and M. E. Charlton. An Investigation of Methods for Visualising Highly Multivariate Datasets. *Case Studies of Visualization in the Social Sciences. Joint Information Systems Committee / ESRC*, pages 55–80, 1998.
- [Hendley *et al.*, 1995] R. J. Hendley, N. S. Drew, A. M. Wood, and R. Beale. Narcissus: Visualising Information. In Nahum D. Gershon and Steve Eick, editors, *Proceedings of the IEEE Symp. Information Visualization (InfoVis '95)*, pages 90–96. IEEE Computer Society Press, 30–31 1995.
- [K. Card, 1999] B. Shneiderman K. Card, J. D. Mackinlay. *Readings in Information Visualization*. Morgan Kaufman, 1999. Chapter 1.
- [Koffka, 1935] Kurt Koffka. *Principles of Gestalt Psychology*. Harcourt-Brace, 1935.
- [Lanzenberger *et al.*, 2003a] Monika Lanzenberger, Silvia Miksch, Susanne Ohmann, and Christian Popow. Applying Information Visualization Techniques to Capture and Explore the Course of Cognitive Behavioral Therapy. In *Proceedings of the 2003 ACM symposium on Applied computing (SAC '03)*, pages 268–274. ACM Press, 2003.
- [Lanzenberger *et al.*, 2003b] Monika Lanzenberger, Silvia Miksch, and Margit Pohl. The Stardines - Visualizing Highly Structured Data. In *Proceedings of the Seventh International Conference on Information Visualization (InfoVis '03)*, pages 47–52. IEEE Computer Society, 2003.
- [Matthews and Roze, 1997] Geoffrey Matthews and Mike Roze. Worm Plots. *Computer Graphics and Applications, IEEE*, 17(6):17–20, 1997.
- [McGinn and Picking, 2003] John McGinn and Richard Picking. The Argument-as Metaphor in Decision-making Visualisation. volume Seventh International Conference on Information Visualization (InfoVis '03), pages 596–599. IEEE Computer Society, 2003.
- [Nakakoji *et al.*, 2001] Kumiyo Nakakoji, Akio Takashima, and Ysuhiko Yamamoto. Cognitive Effects of Animated Visualization in Exploratory Visual Data Analysis. *Fifth International Conference on Information Visualisation, IEEE Computing Society (InfoVis '01)*, pages 77–84, 2001.
- [Noirhomme-Fraiture, 2002] M. Noirhomme-Fraiture. Visualization of Large Data Sets: The Zoom Star Solution. *International Electronic Journal of Symbolic Data Analysis*, 0(0), 2002.
- [Olsen *et al.*, 1993] Kai A. Olsen, Robert R. Korfhage, Kenneth M. Sochats, Michael B. Spring, and James G. Williams. Visualization of a Document Collection: The VIBE System. *Information Processing & Management*, 29(1):69–81, 1993.
- [Rao and Card, 1994] Ramana Rao and Stuart K. Card. The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus+Context Visualization for Tabular Information. In *Proceedings of the ACM Conference Human Factors in Computing Systems (CHI '94)*, pages 318–322. ACM, 1994.
- [Tominski *et al.*, 2003] Christian Tominski, James Abello, and Heidrun Schumann. Interactive Poster: Axes-Based Visualizations for Time Series Data. In *Poster Compendium of IEEE Symposium on Information Visualization (InfoVis '03)*, pages 68–69. IEEE, 2003.
- [Ward, 1994] Matthew O. Ward. XmdvTool: Integrating Multiple Methods for Visualizing Multivariate Data. In *Proceedings of the conference on Visualization (Vis '94)*, pages 326–333. IEEE Computer Society Press, 1994.
- [Ware, 2000] Colin Ware. *Information Visualization*. Morgan Kaufmann Publishers, San Francisco, 2000.