Pre-print version – for citation please use: Smuc, M., Mayr, E., Lammarsch, T., Aigner, W., Miksch, S., & Gärtner, J. (2009). To score or not to score? Tripling insights for participatory design. *IEEE Computer Graphics & Applications*, *29*(3), 29-38.

TO SCORE OR NOT TO SCORE?
HOW TO TRIPLE INSIGHTS FOR PARTICIPATORY DESIGN

Michael Smuc[1]
Eva Mayr[1]
Tim Lammarsch[2]
Wolfgang Aigner[2]
Silvia Miksch[2]
Johannes Gärtner[3]

[1] Research Center KnowComm
Department for Knowledge and Communication Management
Danube University Krems (Austria)

[2] Department of Information and Knowledge Engineering
Danube University Krems (Austria)

[3] XIMES GmbH
Vienna (Austria)

Contact information:
michael.smuc@donau-uni.ac.at
+43 / 2732 / 893 - 2344
eva.mayr@donau-uni.ac.at
+43 / 2732 / 893 – 2320
tim.lammarsch@donau-uni.ac.at
+43 / 2732 / 893 – 2458
wolfgang.aigner@donau-uni.ac.at
+43 / 2732 / 893 – 2452
silvia.miksch@donau-uni.ac.at
+43 / 2732 / 893 – 2450
gaertner@ximes.com
+43 / 1 / 5357920 – 31

Pre-print version – for citation please use: Smuc, M., Mayr, E., Lammarsch, T., Aigner, W., Miksch, S., & Gärtner, J. (2009). To score or not to score? Tripling insights for participatory design. *IEEE Computer Graphics & Applications*, *29*(3), 29-38.

Keywords: Insight, usability engineering, formative evaluation, participatory design

Abstract: Studies recording the number of user insights into data are used in the evaluation and comparison of Visual Analytics tools. However, such insight studies are based on varying definitions of insights, measure different qualitative and quantitative dimensions, and are seldom used during the participatory design phase of Visual Analytics tools. We introduce three levels of insight methodology to be used during the participatory design process and illustrate these using the DisCō project. We started by using conventional "insight counters" which did not provide us with useful results for iterative design, so we went one step further and coded insights in line with the specific purpose of the tools, and successfully gathered information useful for design improvements. Finally, in order to gain an even deeper understanding, we further analyzed relations between insights and prior knowledge by means of a *Relational Insight Organizer* (RIO). The RIO helped characterize *how* users make sense of a tool, as well as where and whether they gain insights. We discuss the potential and prerequisites of these three levels of insight analysis for iterative design.

The Role of the User in Visual Analytics

Visual tools were already used in the ancient world for data analysis (e.g. the abacus for calculation), but human processing abilities and the capacities of these tools restricted analysis to small datasets. It was not until the development of computers with greater processing power that more complex mathematical analyses of huge data sets became possible. In recent years, computers have also been used to develop visual methods and tools which further support the data analysis process. With the advent of the emerging field of "Visual Analytics", the underlying concept of visual tools is taken a step further. In essence, Visual Analytics combines human analytical capabilities with computer processing capacities[1]. In the Human Computer Interaction process new knowledge is generated and "insights" are gained by the user.

One challenge that confronts Visual Analytics is to develop visual analysis tools that best support the user during data analysis to solve the problem at hand – to score. This is frequently achieved by scoring the insights generated using different tools (summative evaluation). Another way to reach this goal is to design novel Visual Analytics tools in a participatory way and to work with users during the design process (formative evaluation). In this paper, we address how "best" to do this and illustrate the issues involved using examples drawn from the DisCō research project.

DisCō – Participatory Design of New Visual Analytics Tools

The aim of the DisCō project [2] is to design novel tools to visually analyze time-oriented data (see Box 1). These tools are developed for use with "Time Intelligence® Solutions" (TIS) software and are targeted at Human Resource planning consultants. At the beginning of the project, the requirements of this target group were assessed by means of a task and user analysis and include optimal staffing or the design of shift work systems. Users have to present their results to their customers and new tools must have an intuitive, easy-to-understand design. DisCō tools are designed with an early focus on users and iterative testing phases.

Insight as an Evaluation Method for Participatory Design

Various methods have been developed in recent years to test the use of visualization tools to support the human reasoning process. Classic benchmarking metrics, such as efficiency and efficacy, proved to be of limited use for either gaining a deeper understanding of the utility of a visualization technique or evaluating the quality of a Visual Analytics tool. Benchmark metrics are typically task based and used in a highly standardized experimental setting, where tasks have to be compact and predefined. The experimental setting forces definitive, unambiguous, and distinctive answers, while time constraints leave little room for deeper elaboration of the findings[3].

Due to the exploratory nature of knowledge discovery in Visual Analytics, new paradigms for testing and evaluation that go "beyond time and errors" were promoted to fill these explanatory gaps[4]. One of these metrics is the qualitative and quantitative measurement of user reported insights[5]. Though this approach was originally developed (and is applied) for summative evaluations[5], it seemed potentially suitable for generating suggestions on how to improve the design of Visual Analytics tools. After all, as insights are assessed in the analysis process, they should also be able to deliver information about the user's knowledge discovery process.

However, one shortcoming of insight methodology is the absence of a widely accepted, clear definition what actually is an insight[5,6]. The fact that "insight" is an everyday word that everybody seems to understand intuitively, yet has a wide variety of different meanings (Wikipedia lists four different meanings) is particularly problematic. In the scientific domain, "insights" have additional meanings, such as the "aha-experience" during problem solving in Psychology or knowledge discovery in Visual Analytics. All of these meanings are in some way related, raising the need for clear differentiation. A clear definition is a prerequisite for valid measurement of insights (see also Box 2). But in the domain of Visual Analytics, defining insights is further complicated by the granularity of the insight gaining process: Is a new insight gained when a pattern is detected in a visualization, when a cognitive script is identified for data analysis, or when a mental model of the whole data set is completed? Different kinds of insights can be of interest in participatory design: how users make sense of a novel Visual Analytics tool at first sight, which discoveries they make in the data, or the mental model they develop about data and data analysis. So, insight studies in participatory design need a pragmatic definition that encompasses all these kinds of insights. Consequently, we have defined an insight as *the understanding gained by an individual using a visualization tool (or parts thereof) for the purpose of data analysis, which is a gradual process towards discovering new knowledge*. By defining insight as a process rather than an outcome, it is clear that insights can only be assessed in the actual process of visual analysis.

Insight Methodology in DisCō

We encountered a number of challenges in using insight methodology (as described by Saraiya et al.[5]) in the DisCō project. First, only a very limited number of expert users actually analyze time related data for HR planning (our application domain), so we were not able to recruit a large enough sample to allow statistical hypothesis testing. Secondly (and most importantly), the resulting "insight counters" were too superficial and did not provide us with detailed suggestions for improving the tools.

Given the small number of experts available, we were only able to interview two appropriate users on their insights into the visualizations. We also interviewed three semi-experts with limited knowledge of the application domain but some background knowledge in data analysis, and compared their results with those of the domain experts. The similarities identified between experts and semi-experts (reported in Smuc et al.[7]) encouraged us to make further use of semi-experts for analysis purposes.

Confronted with the second challenge, we successively refined the methodological approach proposed by Saraiya et al.[5] and adapted it to the needs of our participatory design project. For our formative evaluation, we decided not to train participants on the use of the tools, but to confront them instead with novel tools. This enables us to determine how users make sense of the tools when they first see them[7]. In the rest of the paper we introduce our methodological approach which we hope will be found useful by researchers and practitioners in the Visual Analytics and Usability fields. We discuss the potential and challenges of using insights in the evaluation and participatory design of Visual Analytics tools at three different levels.

Level 1: "In the North" – Counting Insights

Saraiya et al.[5] and North[3] offer a detailed description of a way to analyze insights. Firstly, they ask participants to analyze data using a given tool and to think aloud while doing so. The think-aloud protocols are then coded by domain experts with respect to different characteristics (see Saraiya et al.[5] for a detailed description): the number of insights, the time of the insight, its domain value, its correctness, its directedness, its breadth and depth, and its category (overview, pattern, groups, and details). The values for each of these characteristics can be used to evaluate the tool in questions.

When we examined the coding of these characteristics in detail, we found that "domain value" was coded similarly to "breadth and depth", while "breadth and depth" in return was coded in a similar way to "overview" and "detail". Especially deep insights have a high domain value and focus on details. Despite intensive research, we have so far been unable to identify any studies addressing the redundancies between these insight categories. Correlations between these categories could uncover the redundancies and show whether they measure one or more dimensions. The characteristics also need to be more clearly defined to allow differentiation. But this is beyond the scope of this article.

Applying "Insight Counters" to DisCō

We followed the procedure suggested by Saraiya et al.[5] in our first level of analysis, with the following main results: On average, participants gained more insights into the Multi Scale Plot than the Cycle Plot (see Box 1 for a description of the Visual Analytics tools); both tools supported pattern finding and insights at an overview level, but neither supported any one of the insight categories in particular (overview, patterns, groups, details; see Figure 1).

Figure 1. Overview of evaluation results at level 1 for all participants

We found these results interesting when comparing the two new tools and the different users groups (experts and semi-experts). However, we encountered problems in suggesting improvements for the novel Visual Analytics tools, because no salient insight profile emerged. Therefore, we decided to look more closely at the insights in order to determine what the study participants contribute when it comes to improving the Visual Analytics tools.

Level 2: "Off to New Horizons" – Categorizing Insights from the Bottom-Up and the Top-Down

Apart from the categories introduced by Saraiya et al.[5], what other options are available for coding insights? North suggests generating these categories from the *bottom-up*, that is, from the insights themselves by means of content analysis: By searching for similar insights and naming the resulting clusters according to their meaning, categories can be generated from the think-aloud-protocols.

A different approach that fits well to a participatory design process is to define categories in advance. This *top-down* process can be aligned with the intended purpose of a tool: Most Visual Analytics tools are designed to support specific kinds of analytic processes. Therefore, insight categories can be defined a priori for each of these analytic processes. If the users generate many insights related to its intended purpose, the tool is helpful to the user; if there are only a few or no such insights, the tool fails to accomplish its intended purpose and its design needs to be reconsidered. An analysis at this level requires a tight integration between the tool's development, design, and evaluation processes: to define top-down insight categories, interpret results, and improve the tool accordingly.

Applying "Bottom-Up and Top-Down Insight Categories" to DisCō

Using the *bottom-up* approach in DisCō, we found that our participants reported two kinds of insights (see Table 1): insights into the data, but also insights into the tool (omitted from the results presented at level 1). These tool insights encompass the users' understanding of how the tool works, how it should be read, and where further improvements are needed to better support data analysis. Our analyses of data and tool insights for the two tools showed that although more data insights were gained for the Multi Scale Plot, nearly the same number of tool insights was generated for both plots (see Figure 2).

Table 1. Examples for each category generated by a bottom-up and a top-down approach.

Bottom-Up		
Data insight		"It decreases until 6 a.m. when it reaches a minimum. I assume this is due to […], to my knowledge, the change of shift." (Multi Scale Plot)
Tool insight		"The more green, the fewer assignments; the more blue, the more assignments." (Multi Scale Plot)
Top-Down		
Multi Scale Plot	Overview	"On average, Sundays are rather low."
	Detail	"It peaks at noon. It's always darkest then."
Cycle Plot	Cycle	"Starting in the morning, it rises to a peak around 10 or 11 a.m. It then calms down by noon, but there is a second peak around 4 or 5 p.m., after which it decreases again."
	Trend	"The first Monday is high, the second is lower, but it rises again on the third and fourth."

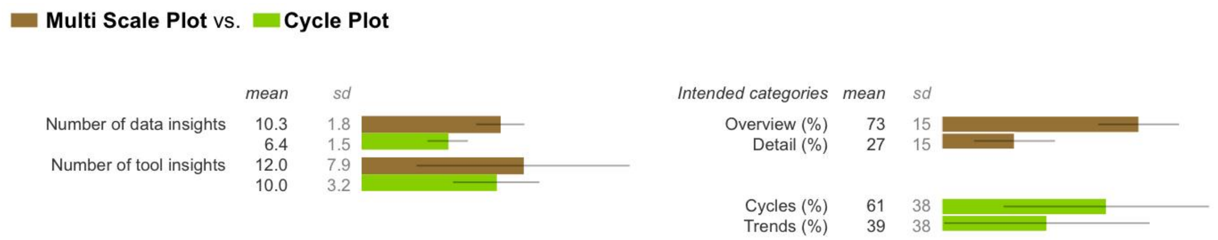

Figure 2. Overview of evaluation results at level 2: bottom-up (left) and top down (right).

Beyond counting them, a qualitative analysis of tool insights is interesting as it shows how users without any prior knowledge make sense of a novel tool. For example, all participants in our study recognized the "calendar metaphor" ad hoc for the Multi Scale Plot. This metaphor helped them immensely in understanding how the tool worked. As a consequence, the tool could be renamed to reflect this metaphor and other tools could be designed and built using similar, well-known, easy-to-apply metaphors like the calendar. This applies in particular to the Cycle Plot, since most participants had problems in understanding how it worked.

Taking the *top-down* approach, we generated insight categories in advance based on the analytic processes each tool was designed to support (see Box 1). Since the intention of the Cycle Plot is to help users to identify trends and cycles, we coded data insights for this plot as either cycle or trend data insights (see Table 1). The Multi Scale Plot is designed to help users gain insights at both the overview and the detail levels. We again coded data insights accordingly as overview or detail data insights (see Table 1 and Figure 2). At first sight, these categories are similar to those defined by Saraiya et al.[5]. However, here we also coded both for patterns and groups as to whether they focused on an overview or a detail level. In contrast to the results obtained at level 1, more details were categorized at this level. One possible explanation for this could be that patterns and groups are coded preferentially as a detail level. For the new categories "trends" and "cycles", a clear picture emerged: Cyclic structures led more often to insights. However, it should also be noted that the standard deviation for these two categories is rather high.

These findings reveal the need for improvement for both plots: Neither supports the user in generating insights into both categories to an equivalent extent (as originally intended). A possible improvement could be to allow the user to switch interactively between trend and cycle lines or overview and detailed view respectively.

More detailed analyses of the transcripts from level 2 indicated that tool insights often precede any related data insights. As a consequence, we looked more closely at the connections between insights, that is, how they build on one another and how a complete picture is obtained.

Level 3: "Arriving in RIO" – the Relational Insight Organizer

In the prior two levels, we only coded characteristics of single insights. But insights always build on prior knowledge and insights into the data and the tool. That is why most evaluations only focus on expert users or imply intensive a priori training. To better understand how such expert users gain insights when they use Visual Analytics tools, an analysis of the insight generation process and how insights build on each other is required.

How can these relations between insights be best analyzed? In computer-supported collaborative learning, Suthers et al. suggest an "uptake graph" showing how the utterances made by learners build on their own prior utterances and those of other learners on a timeline[8]. We applied this approach to the insight timeline of one participant in our study and drew a "relational insight organizer" (RIO, see Figure 3). The RIO consists of multiple rows of icons along a timeline. The upper row shows prior knowledge, the next rows each show one (bottom-up or top-down) insight category (in our case, insights into both the tool and the data). The icons are positioned according to when they were first mentioned. For each insight identified in the think-aloud-protocol, note is made of whether it required prior knowledge and whether it built on a prior insight. To ensure the reliability of these relations, we found it beneficial to have them rated by at least two researchers. When the insight is built on prior knowledge, a prior knowledge icon is added at this time point in the first row and an arrow is drawn to indicate the relation. Similarly, for insights that build on prior insights, an arrow is drawn from the first to the second insight. To further aid interpretation of the relations between utterances, Suthers adds content (parts of the utterances) to the graph. In RIO, we followed this suggestion to a certain degree for easier interpretation, naming insights according to their assigned category (see level 2). For some analyses, it might be interesting to add annotations, parts of the transcripts, or screenshots to help other researchers and designers reconstruct the interpretations. An example is given in Figure 3 for user 3. (See also Box 2).

Applying RIO to DisCō

When we used this type of visualization for the insights into both novel Visual Analytics tools in DisCō, we found the visual analysis of the insights of our participants particularly enlightening. In a first analysis step, we took a process-oriented view of success/failure stories for individual users to identify factors that allowed them to score or not. In a second step, these individual processes could be compared to identify more general patterns of analysis. These patterns could then be used to align tool design to user analysis processes.

In Figure 3, we show two RIOs for the Cycle Plot. The upper RIO shows insight generation by a semi-expert. Relatively quickly, this user had the insight that the tool shows trends over a four-week period (H1). The user's data insights prior to the explanation of the tool can be divided into three phases: The first three data insights focus on the daily cycle (C1 to C3), the next three are generated with respect to trends (T4 to T6), and his final insight identifies an interaction between cycles and trends (CT7). Interestingly, we found that these first two phases were common to nearly all other participants: They focus initially on the daily cycle, then change to trends. Only the expert user (shown in Figure 3 in the lower RIO) skipped the first phase and primarily analyzed trends, mentioning only at the end that "a daily cycle can also be seen". This result allows us to draw two inferences: Firstly, users seem to find it difficult to switch quickly and frequently between interpreting cycles and trends (which they all did for the Multi Scale Plot's intended overview and detail categories) and, secondly, the salience of cycles seems to be higher because corresponding insights are nearly always generated first and (as can be seen in the RIO for user 1) they do not require a related preceding tool insight.

Based on these results, we suggested two improvements to the tool: an increase in the salience of the trends and the provision of a switch between the cycle-salient and trend-salient view.

A second pattern identified in the RIOs displayed in Figure 3 is the use of prior knowledge by experts in comparison to semi-experts: User 3 (an expert) used more prior knowledge than user 1 (a semi-expert) to understand how the tool works and how it should be used for data insight generation. This pattern was also observed for the second tool and the second expert. This indicates that an expert's domain knowledge guides the use of a Visual Analytics tool to a great extent, but does not have as much of an influence on their interpretation of data. This might be a problem specific to the HR planning domain, because customers come from different sectors of industry, requiring less sector-specific knowledge than analysis-specific knowledge. The expert's cognitive scripts and knowledge of analyzing time-oriented HR data greatly influence how they extract information from a tool.

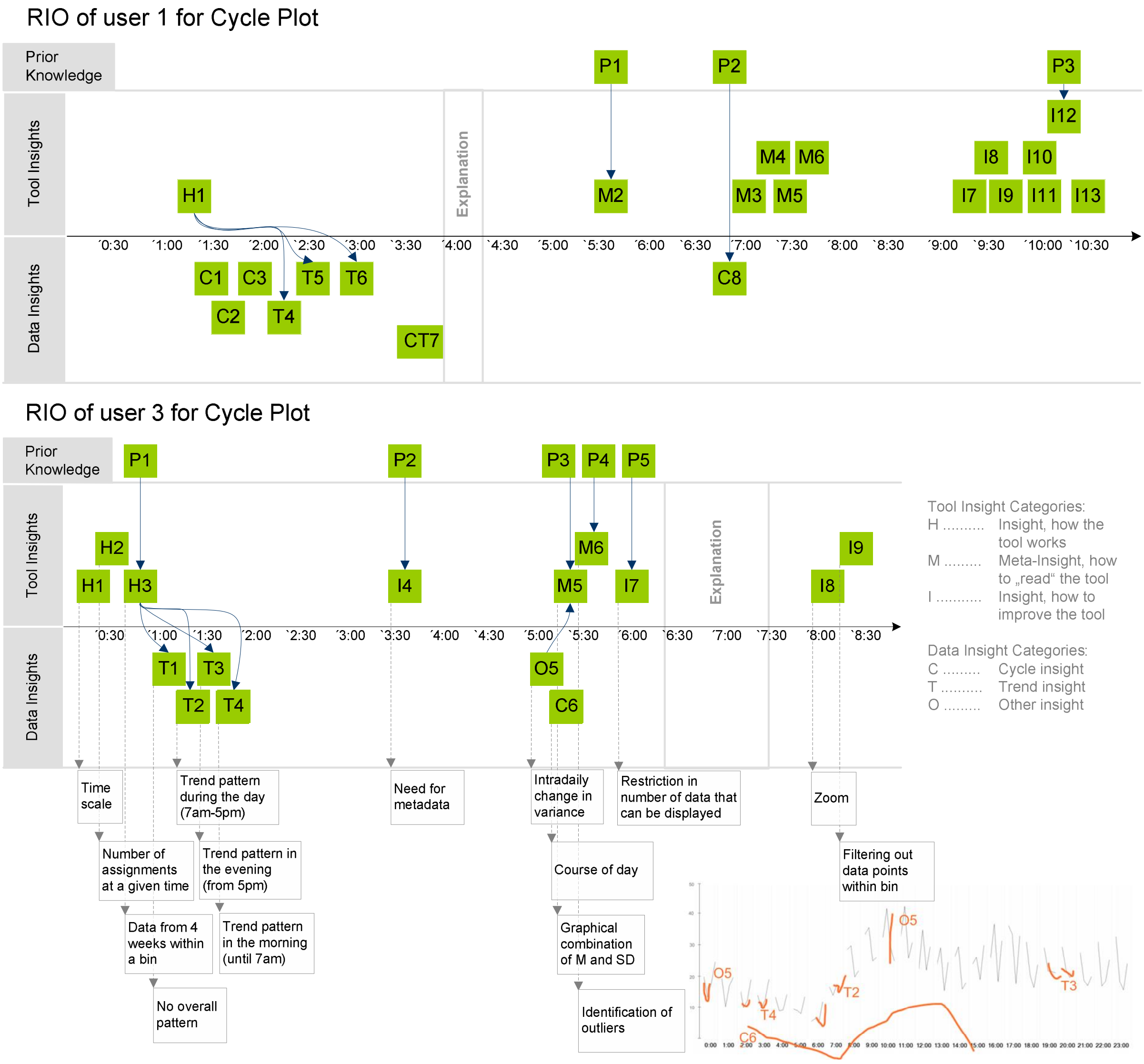

Figure 3. Relational Insight Organizers (RIO) of user 1 (top, a semi-expert) and user 3 (bottom, an expert) for the Cycle Plot.

Moving up North or to RIO? A Guidepost

We introduced three different levels of analysis that can be applied to insights assessed via think-aloud-protocols: (1) insight counters, (2) bottom-up and top-down insight categories, and (3) the Relational Insight Organizer, RIO. But which of these levels is the most appropriate? In our opinion, no single approach is superior. When applying the various levels in DisCō, we always gained important results. Each level provides different answers to different questions. Our methodological exploration leads to the conclusion that the appropriate level of analysis depends primarily on the evaluation goals (e.g. whether it is applied during summative evaluation or iterative testing or how much time is available for analysis) and the research question (e.g. whether tools are to be compared or improved and the kind of results needed). In the following, we provide assistance with choosing the route to take, that is, to go North or to go to RIO. Table 2 provides an overview of all levels.

Table 2. Comparison of the three levels of insight analysis with respect to the outcomes, answerable research questions, potential, problems, and possible design improvements.

Level	**Research Questions**	**Output**	**Potential / Benefits**	**Challenges / Problems**	**Design Questions**
1	Comparison of tools, users or user groups	Number of insights, timeline, insight categories, correctness	Easy to apply, easy to interpret, quantitative results	No qualitative findings	Which of the different visualizations should be selected?
2	What kinds of insights does a tool promote?	Intended insight categories, tool insights	Findings tailored to the Visual Analytics tool	More laborious data analysis, only some qualitative findings	How can the tool be improved to support the intended analysis processes?
3	How does the tool promote the generation of insights?	Relations between insights, analysis process	More qualitative findings related to the process, shows relations between insights instead of simply counting	Very laborious data analysis, only possible with small samples	Where did users fail / succeed in gaining the intended insights? How can the process be supported?

When we look again at *level 1*, one of its major advantages is that the analysis steps are relatively well-defined (with the exception of the above-mentioned difficulties with some of the categories) and can be applied with the least effort. However, this analysis level is also time consuming. We found it most appropriate when used to analyze differences between groups of users or visualizations. Saraiya et al.[5] also used "insight counters" to compare different Visual Analytics tools with regard to their ability to support user knowledge discovery processes. One restriction of insight counters is that they provide more quantitative than qualitative results. This is beneficial insofar as they are easy to interpret, but also problematic since no qualitative findings for design improvements are gained. In participatory design, this level is only beneficial if we have a direct comparison between alternative tools or different variants of the same tool.

At *level 2*, top-down or bottom-up strategies can be applied to generate new, tailored insight categories. Top-down strategies are particularly important for insight studies in participatory design: By defining the kind of insights the novel Visual Analytics tool should promote in advance, the results of the user study demonstrate whether or not the tool fulfils its intended purpose. This can be less time-consuming than a full analysis at level 1 (although one should consider using both levels), but the output can be much more relevant for design improvements.

Bottom-up strategies are more time-consuming in analysis because clustering and content analysis have to be carried out before the insights can be coded. But a bottom-up approach is often worth the additional effort as it demonstrates those insights promoted by the novel tool that the developers and researchers did not think of in advance. In our study, we found that users gain both tool and data insights, which may have been provoked by the omission of the a priori training phase in the DisCō project. But these insights allow us to observe how participants gain an understanding of how a novel Visual Analytics tool works. Although this is rare for insight studies to date, we would like to encourage this approach in other participatory design studies as it provides valuable findings on how users make sense of a tool, how they struggle with visualization elements, how they learn to “read” the visualizations, and how they start to work with the tool in the analysis processes.

Analyses at level 2 can be more laborious than at level 1, especially when insight categories are generated bottom-up. But both bottom-up and top-down categories provide useful information on the kind of insights the tool promotes. This information can be used to further improve the tool in the next iterative design phase.

Level 3 contains the most time-consuming insight analyses: data and tool insights are visually plotted on the user’s analysis process timeline. In the RIO, insights are related to the prior insights and to the prior knowledge on which they are built and to the subsequent insights that make use of them. This analysis has to be conducted separately for each individual user and may, therefore, be restricted to sub-samples only. In our experience, to provide visual clarity a RIO for a think-aloud-protocol of approximately one hour can be drawn on one sheet of paper.

One benefit of RIO over the first two levels is that we remain close to the visual analysis process instead of being restricted to summative insight counting. As with Visual Analytics, we move from pure computational data analysis to visual data analysis. The latter makes it easier to obtain a more complete picture, to see relations, and to interpret data more qualitatively. By taking this step, we not only assess our measures during the analysis process, we also interpret them in a process-like manner. This aids the design of Visual Analytics tools enormously: We can see where a user is able to generate insight and where a user succeeds or fails to gain the intended insight, that is, where a user does score or not. This process view helps to improve the design of Visual Analytics tools throughout the data analysis process.

DisCō’s Travelogue

In an early, iterative usability engineering phase, we conducted analysis using all three levels described above to test the insight methodology. RIO turned out to be valuable in providing insights leading to design improvements. One restraint of our study is that RIO was not tested using interactive elements, which might have lead to additional errors and detours. At this early stage in DisCō, we used only mock-ups of the tools, but did gain insight-based suggestions for where interactions were necessary. These findings were subsequently used to develop the Visual Analytics tools further.

Our plan for the next iterations of formative testing is to use RIO for single, early stage developments, but analyses at levels 1 and 2 for multiple comparisons and summative evaluation at the end of the project. In other words, we are already heading back North, but are prepared to come back for a "samba in RIO".
One limitation of all the above levels of analysis is that collecting and analyzing insights is time consuming and should be used only when necessary. Insight analyses only proved appropriate for some research questions (see Table 2) and can be combined or substituted with other less time-consuming methodologies. For example, in DisCō we also relied on questionnaires and traditional post tests. Nonetheless, insights provide a good illustration of users' understanding gained when using Visual Analytics tools.

To conclude our journey from North to RIO, we suggest using all three levels of analysis at different phases of participatory design: level 1 for comparison of different tools, tool variants, or groups of users; level 2 for testing the goodness of fit of a designer's intentions; and level 3 for better aligning the tool with the user insight generation process. We recommend starting the journey in the North and only travelling to new horizons if the results at this stage lack salience. If user insights do not meet the tool's purpose on this trip, feel free to put on your dancing shoes and join in the samba in RIO.

Acknowledgements
The DisCō project (project number: 813388) is supported by the "FIT-IT Visual Computing" program of the Federal Ministry of Transport, Innovation and Technology in Austria. We would like to thank our colleagues Hanna Risku, Karin Siebenhandl, Kay Muehlmann, and Florian Windhager as well as the other members of the DisCō project consortium for their feedback on earlier versions of this paper.

Box 1: **Novel Visual Analytics tools developed in DisCō**
Analysis of time-oriented data includes exploration of trends, patterns, and relationships of multivariate information. Time data is difficult to analyze because of its complex, natural (e.g. seasons, days) and social (e.g. business years, holidays) structure.[9]
In the case study presented here as an example, two Visual Analytics tools for time-oriented data are analyzed: the Cycle Plot[10] and the Multi Scale Plot[11]. The aim of the Cycle Plot (see Figure A) is to help the user differentiate trend and cyclic structures in data. The aim of the Multi Scale Plot is to show as much data as possible in a limited space (see Figure B). It makes use of the structure of time and was modified to provide insight into overview and detail in a single visualization.

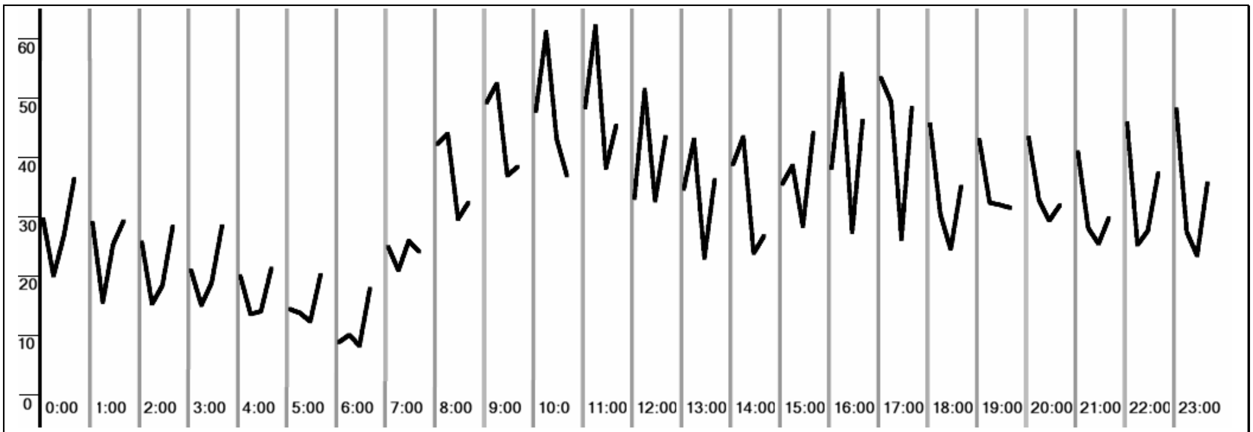

Figure A. Example of a Cycle Plot showing the number of police assignment in the Monday daily cycle (24 h, x-axis). Within each hour, data from 4 successive Mondays is displayed to show trends.

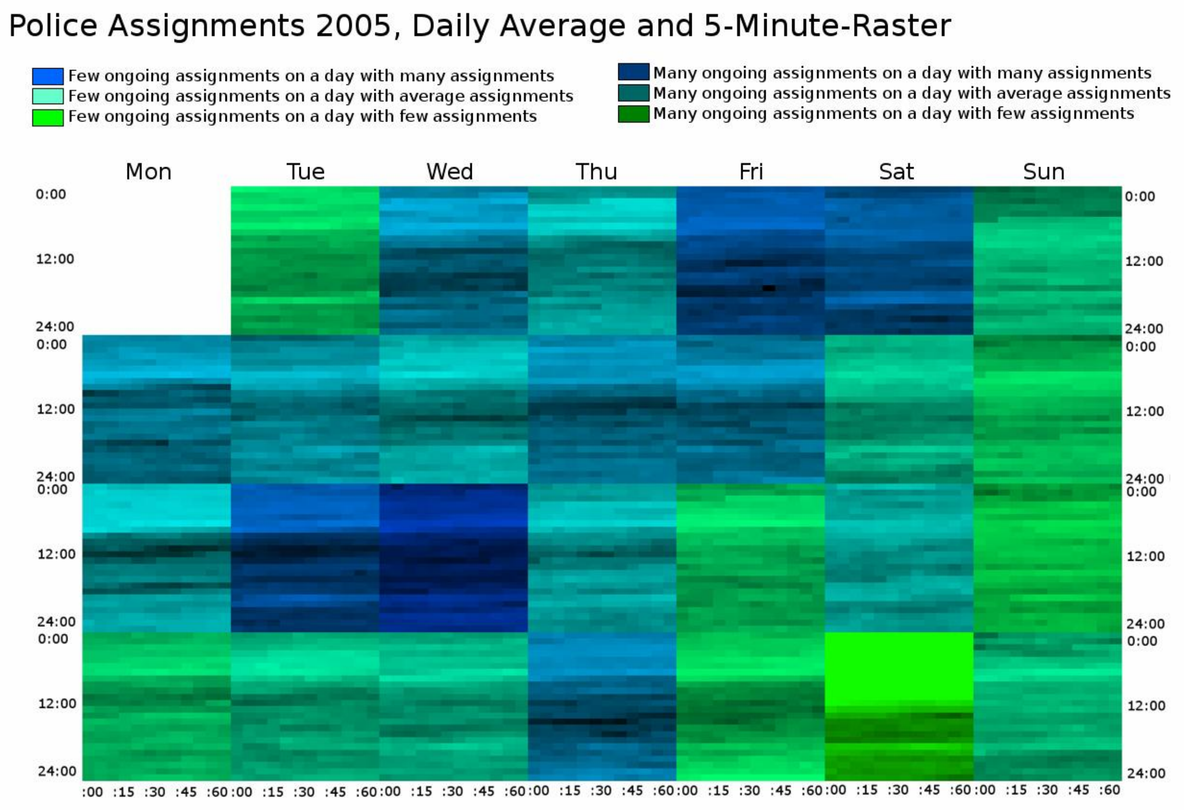

Figure B. Example of a Multi Scale Plot. Like a calendar, each block shows one day. Inside each block, the rows represent hours and each pixel in a row represents a 5-minute-interval. In this visualization, 7803 data points are displayed.

Box 2: **Moving into Unchartered Waters – What's Next on Our Journey into Understanding Insights?**

Until now, insights in Visual Analytics were defined rather narrowly as knowledge discoveries regarding the data. In our study, we categorized user insights from the bottom-up and found that they also frequently included insights into the tool. Such tool insights offer valuable information for the design of Visual Analytics tools.

Further work on insight methodology should address how interaction with the tool, the data and the annotations relates to user insights. The amount of elaboration is also an interesting factor (one user in our study took several minutes to elaborate on one tool insight). With some refinements, both ideas could be visualized in RIO (as new relations and duration-dependent insight blocks).

An important contribution for the community and for understanding user insight generation would also be a cognitive theory of the user in visual data analysis. Important sources for such a theory could be findings from psychology and sensemaking, with the inclusion of the elaboration cycles introduced by sensemaking frameworks in particular worth consideration.[12]

---

[1] J.J. Thomas and K.A. Cook, *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, IEEE CS Press, http://nvac.pnl.gov/docs/RD_Agenda_VisualAnalytics.pdf (2005).

[2] from Latin "discō" (inf. discere) meaning "I learn"; http://www.donau-uni.ac.at/disco

[3] C. North, "Toward Measuring Visualization Insight," *IEEE Comput. Graph*., vol. 26, no. 3, 2005, pp. 6-9.

[4] see the BELIV'06 (http://www.dis.uniroma1.it/~beliv06/) and BELIV'08 (http://www.dis.uniroma1.it/~beliv08/) Workshops and C. Plaisant, J. Fekete, and G. Grinstein, "Promoting Insight-Based Evaluation of Visualizations: From Contest to Benchmark Repository," *IEEE Trans. Vis. Comput. Graph.*, vol. 14, no. 1, 2008, pp. 120-134.

[5] P.B. Saraiya, C. North, and K. Duca, "An Insight-Based Methodology for Evaluating Bioinformatics Visualizations," *IEEE Trans. Vis. Comput. Graph*., vol. 11, no. 4, 2005, pp. 443-456.

[6] G. W. Hartmann, "The concept and criteria of insight," *Psychol. Rev.,* vol. 38, 1931, pp. 242–252; G. Jones, "Testing Two Cognitive Theories of Insight," *J. Exp. Psychol. Learn. Mem. Cogn*., vol. 29, 2003, pp. 1017–1027.

[7] M. Smuc et. al., "Visualizations at First Sight. Do Insights Require Training?" *Proceedings of the USAB 2008 - Usability & HCI for Education and Work*, Springer, Lecture Notes in Computer Science (LNCS), 2008.

8 D. Suthers et al., "An Abstract Transcript Notation for Analyzing Interactional Construction of Meaning in Online Learning," *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on the System Sciences*, 2007, p. 4c.

[9] W. Aigner et al., "Visualizing Time-Oriented Data – A Systematic View," *Computers & Graphics*, vol. 31, no. 3, 2007, pp. 401-409.

[10] W. Cleveland, *Visualizing Data*. Hobart Press, 1993.

[11] M. Shimabukuro et al., "Coordinated Views to Assist Exploration of Spatio-Temporal Data: A Case Study," *Proceedings of the 2nd International Conference on Coordinated and Multiple Views in Exploratory Visualization* (CMV04), IEEE CS Press, 2004, pp. 107–117.

[12] P. Pirolli and S. Card, "The Sensemaking Process and Leverage Points for Analyst Technology as Identified through Cognitive Task Analysis". *Proceedings of 2005 International Conference on Intelligence Analysis*, McLean, 2005, pp. 2-4.