Interactive analysis of Likert scale data using a multichart visualization tool

Conference Paper · October 2011

4 authors:

Fabio Petrillo
University of Québec in Chicoutimi

André S Spritzer
Universidade Federal do Rio Grande do Sul

Carla Dal Sasso Freitas
Universidade Federal do Rio Grande do Sul

Marcelo Pimenta
Universidade Federal do Rio Grande do Sul

Interactive Analysis of Likert Scale Data using a Multichart Visualization Tool

Fabio Petrillo, André Suslik Spritzer, Carla Dal Sasso Freitas, Marcelo Pimenta
Institute of Informatics - Federal University of Rio Grande do Sul (UFRGS)
Porto Alegre, Brazil
fabio@petrillo.com, {spritzer, carla, pimenta}@inf.ufrgs.br

ABSTRACT

Likert scale data correspond to data collected from questionnaires, where there are several statements about a fact, situation, design or technique, and the respondents must specify their level of agreement with each statement by marking one of several ordered alternatives. Often, five choices are posed to respondents (strongly agree, agree, neither agree nor disagree, disagree and strongly disagree) for each statement. Data obtained with such surveys are translated to numerical scores to undergo statistical analysis, and are presented graphically in a number of different ways. When evaluating information visualization techniques, for comparison purposes it is often necessary to apply the same Likert-scale based questionnaire to different views or techniques. Most of the existing tools, however, do not support the simultaneous analysis of multiple Likert scale datasets. This paper presents an interactive visualization technique to help the analysis of Likert scale data obtained from several, alternative sources. The technique is based on a hybrid table and bar chart view, and provides different interactive tools to support data exploration. We describe the use of our tool in a case study where we compare different graph layouts.

Categories and Subject Descriptors

H.5.2 [**Information Systems Applications**]: User Interfaces – Evaluation/methodology; D.2.2 [**Software Engineering**]: Design Tools and Techniques – User Interfaces

Keywords

Information Visualization, User Interfaces, Likert scale

1. INTRODUCTION

One of the most used methods for collecting data during the evaluation of user interfaces or interactive applications is applying a questionnaire that gathers users' opinions after they use the interface or the application for some time. Whether the selected evaluation model is a simple, free exploration of the software, or a formal experiment with a strict protocol, questionnaires allow one to collect information about the users' subjective satisfaction, and as such can be difficult to analyze.

To facilitate the interpretation of users' responses to questionnaires, one can use the Likert scale to build the set of possible answers. Originally introduced by psychologist Rensis Likert in 1932 [10], the Likert scale has become the most widely used psychometric response scale. The scale consists of several statements, or Likert items, to which respondents specify their level of agreement by marking, for each statement, one of several ordered response alternatives. While most often used with five response alternatives, this is not a rule, with seven and ten-point scales also being common. A typical five-point Likert scale consists of the following ordered response options: 1 - Strongly Disagree, 2 - Disagree, 3 - Undecided, 4 - Agree, and 5 - Strongly Agree. Sometimes the neutral response is omitted, forcing respondents to take a position of either agreement or disagreement. There are also variations to the ordered response options utilized, which could range for instance, from Very Not Effective to Very Effective, Never to Very Frequently, Unimportant to Important, and so on.

Analyses of responses to such questionnaires usually aim at identifying pros and cons of the evaluated user interface, detecting patterns of responses or correlation between answers to different statements. Depending on the number of statements in the questionnaire and the length of the Likert scale, results can be difficult to analyze. The most common way of analyzing the responses of such questionnaires is to build static line or bar charts (see, for example, Figures 1 to 3), and compute some basic statistics.

In information visualization, when evaluating a technique it is very often necessary to compare different alternative designs, which imply in comparing responses to different questionnaires, and consequently comparing several charts displaying several Likert-scale data sets.

In the present paper we introduce a visualization-based tool for the analysis of several Likert-scale data sets. The tool is based on displaying multiple bar charts, and allows interacting with the data through filtering, selection and manipulation, so the designer can reach a more solid interpretation of results within a shorter timeframe.

The remainder of the text is organized as follows. Next section presents the background and a review of related work. Section 3 describes MCLikertVis, our tool for supporting the analysis and interpretation of multiple Likert-scale data sets while Section 4 presents a case study to emphasize the contributions of the work. Finally, in Section 5 we draw some conclusions and comment on future research.

2. BACKGROUND AND RELATED WORK

2.1 Analyzing Likert Scale Data

Likert scales are easy to build and analyze, but are also subject to distortions. Among other reasons, these can be caused by respondents avoiding extremes, tending to agree with every statement, or not answering entirely truthfully in order to portray themselves in a certain way. Problems can also arise from badly designed questions that may lead to meaningless answers.

To analyze Likert scale data, one can summate responses to obtain an overall score for a group of related items, or analyze each item individually. During the analysis, the discrete nature of item responses should be respected, since ignoring this can lead to inferential errors. While the responses are indeed ordinal, the numbers used to represent them are just a coding. To summarize the data, therefore, one should avoid tools that are geared at continuous values, such as means, because they would have no meaning within this context (e.g. there is no mean between the concepts of Agree and Undecided). A better approach would be to use counts or percentages occurring in the various response categories to see answer distribution, along with the median or the mode and the interquartile range, which shows the variability of the responses in the middle 50.

After descriptive analysis is done, one can make use of inferential techniques to test hypotheses and compare how different groups of subjects responded to certain statements. These techniques include analysis of variance techniques such as Mann-Whitney (to compare unrelated samples), Wilcoxon Signed Rank (for paired samples), and Kuskal-Wallis (for three or more samples). Further summarizing the data by combining response categories into two nominal categories (i.e. summing Strongly Agree and Agree into an Agree category, and Strongly Disagree and Disagree into a Disagree category) allows for more analytical options, such as the Chi-square test.

It is important to bear in mind that there is no right or wrong way of analyzing Likert data. It really depends on the problem at hand. While for some problems a certain technique might yield meaningful answers to a meaningful question, the same technique can be completely useless in other contexts.

While Likert-like surveys can be created by hand and analyzed by hand or with the help of spreadsheet or statistics software, there are some online tools that help make the whole process more automatic, with features to create and distribute the surveys, and analyze and visualize the results. Amongst these tools, we find the commercial tools WAMMI [4], SurveyGizmo [2], and SuperSurvey [1], with the latter two also having free versions. Another free alternative for simple survey building is Google Docs.

2.2 Visual Reporting on Likert Scale Data

There is no standard visual representation for Likert data, with the choice of which one to use depending highly on what is being studied and how the analyst wants to show it. The most widely used chart type for Likert scales is the **bar chart** (Figure 1), which places the count or percentage in one axis (usually the vertical) and the ordered response categories on the other (the horizontal axis). Response distribution for each category is represented by a bar with height corresponding to the amount or percentage of respondents that chose it for a given Likert item (statement). For comparison purposes, it is sometimes interesting to show several Likert items in the same bar chart by placing different-colored bars side by side, with each color corresponding to a different statement (Figure 2).

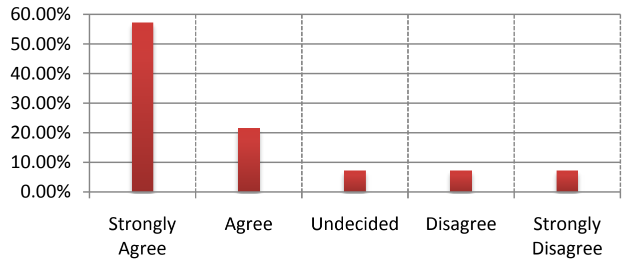

Figure 1: Simple bar chart.

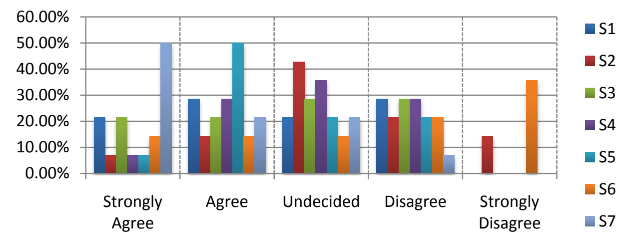

Figure 2: Multi-statement bar chart.

Another approach for showing multiple statements simultaneously is to use **stacked bar charts** (Figure 3), which show response distribution by subdividing, for each statement, a single bar into several colors, each representing a response category. The colored subdivisions follow the same order in every bar, respecting the ordinal trait of the categories, and have size proportional to the amount of respondents that chose the category they represent.

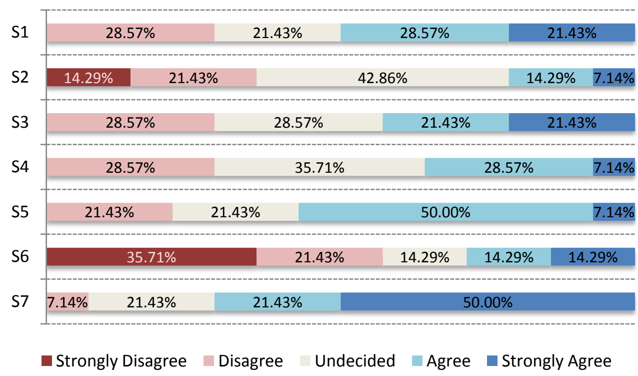

Figure 3: Stacked bar chart.

A variation of stacked bar charts is the **centered stacked bar chart** (Figure 4). It works in a similar fashion to a typical bar chart, but removes the neutral response category in order to place the right end of the mildest negative category subdivision at the same position of the horizontal axis as the left end of the mildest positive category subdivision. This creates a central line dividing positive from negative responses, allowing for the skew between them to be seen more easily.

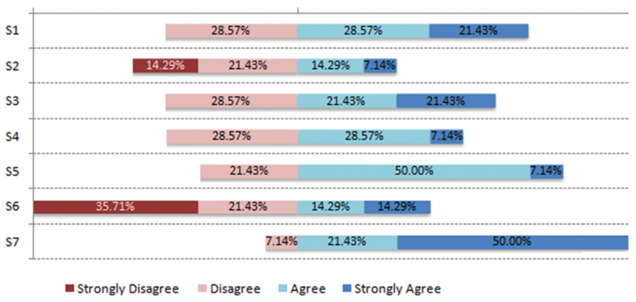

Figure 4: Centered stacked bar chart.

Other sometimes used approaches to visualize Likert data include **dot charts**, which substitute dots for bars, placing them where the upper end of a bar would be, and **pie charts**, which for each statement creates colored subdivisions in a circle corresponding to each response category, with each subdivision assuming a size proportional to the percentage of respondents that chose its category.

Different computational tools can be used to build charts for Likert data. These range from generic spreadsheet applications, such as Microsoft Excel, to statistics tools such as SPSS and the R programming language. Particularly interesting is Tableau [3], a commercial MS Windows-based software package that can be used to build rich, interactive visualizations of tabular data (including, or course, Likert scales). A free version of Tableau, called Tableau Public, is available.

3. MULTICHART LIKERT SCALE VISUALIZATION

Despite the existence of interactive analysis tools such as Tableau [3], the common way of presenting Likert scale data is with static tables and simple charts. This imposes difficulties, especially when one needs to compare multiple datasets, i.e., responses obtained from several instances of the same questionnaire (the same set of Likert items), each evaluating a different design. This would be the case, for example, of comparing different visualization tools that target the same problem, or various designs of a new cell phone. In this case, each visualization tool or alternative design would be evaluated through the same questionnaire resulting different datasets.

Our proposal to address this issue is by means of a new visualization tool called Multichart Likert Scale Visualization, or MCLikertVis, which consists of a matrix of charts (Stacked Bar, Simple Bars and Pie) organized so that the lines represent datasets, while the columns are the Likert items (Figure 5). This visualization was inspired by ScatterDice [7], in which the user can navigate in a scatterplot matrix and interact with the individual plots.

Selecting a chart in this matrix (see area 1, Figure 6) triggers its display in the details panel (see area 2, Figure 6). Below this panel, basic statistics such as Mode, 1st Quartile, Median (2nd Quartile), 3rd Quartile and Interquartile Range are shown (area 3, Figure 6). Area 4, also in Figure 6, displays a set of widgets that allows users to simplify the charts by aggregating the responses "Agree" and "Disagree", and to apply different filters that exclude from the matrix those charts that do not comply with the filtering parameters. Users can also select the type of chart to be displayed in the matrix, with the current options being Stacked Bar, Simple Bar and Pie charts.

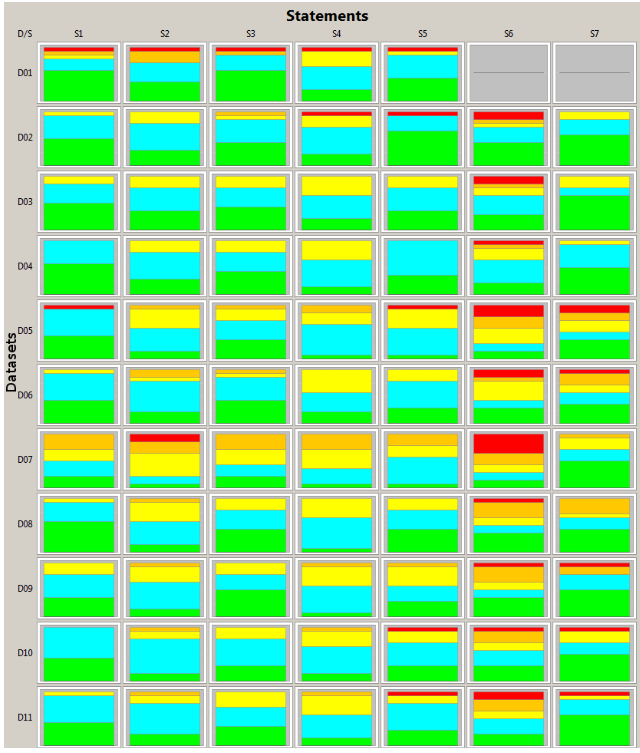

Figure 5: MCLikertVis stacked bar chart default visualization.

Figure 7 shows an example where the user has aggregated the responses "Agree" and "Disagree", so that the charts have only three possible levels. Also, we can observe that the matrix of charts is sparse because the user has set a filter of "disagree level" at 20%. If the datasets correspond to different design alternatives for a web site, for example, this filtering allows the user to quickly analyze the alternatives that correspond to good or bad designs.

Figures 8 and 9 show other possible visualizations with MCLikertVis, based on Simple Bar and Pie charts, respectively. Both examples were obtained with the application of filters. One can also observe that the application of filters are actually queries over the datasets, with the results being shown in different panes (notice the TabbedPane at the top of main area). By browsing through the different panes, the user can rapidly compare several filtered views of the same data.

Additionally, the tool provides the ability of exporting graphical output in PNG, JPEG or GIF, and printing the results.

The process of building such visualizations begins with lo-

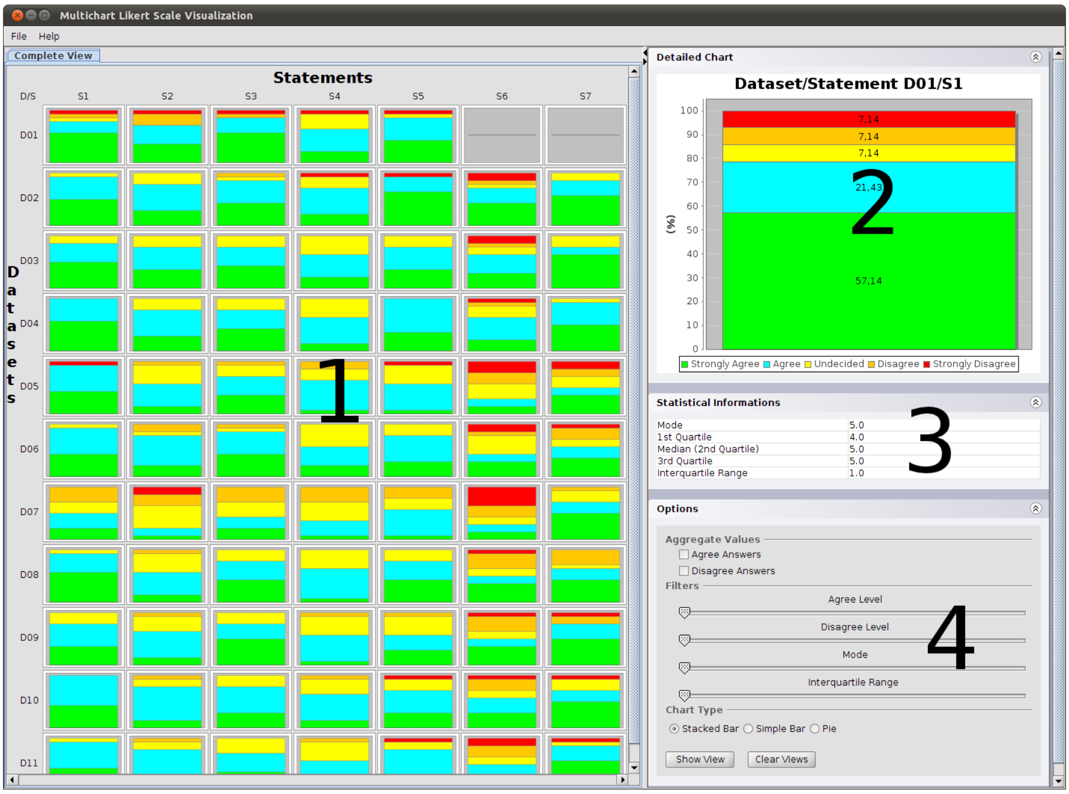

Figure 6: MCLikertVis - Standard View.

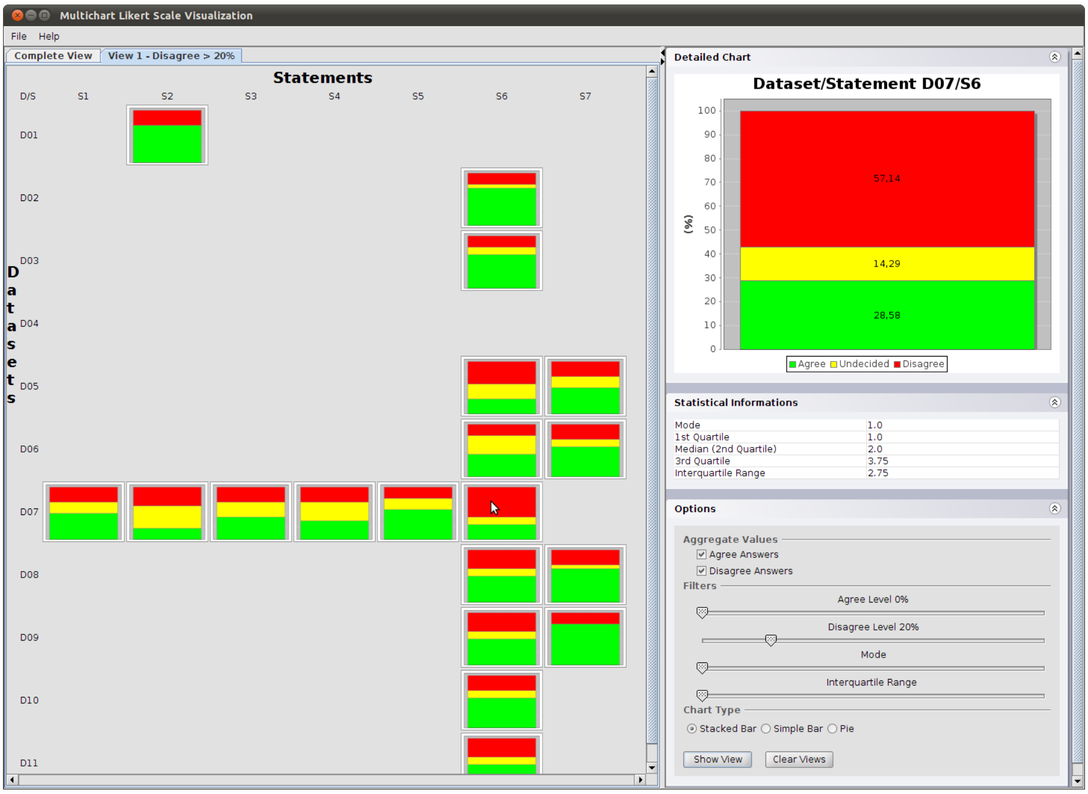

Figure 7: MCLikertVis - filtered and chart details of D07/S6.

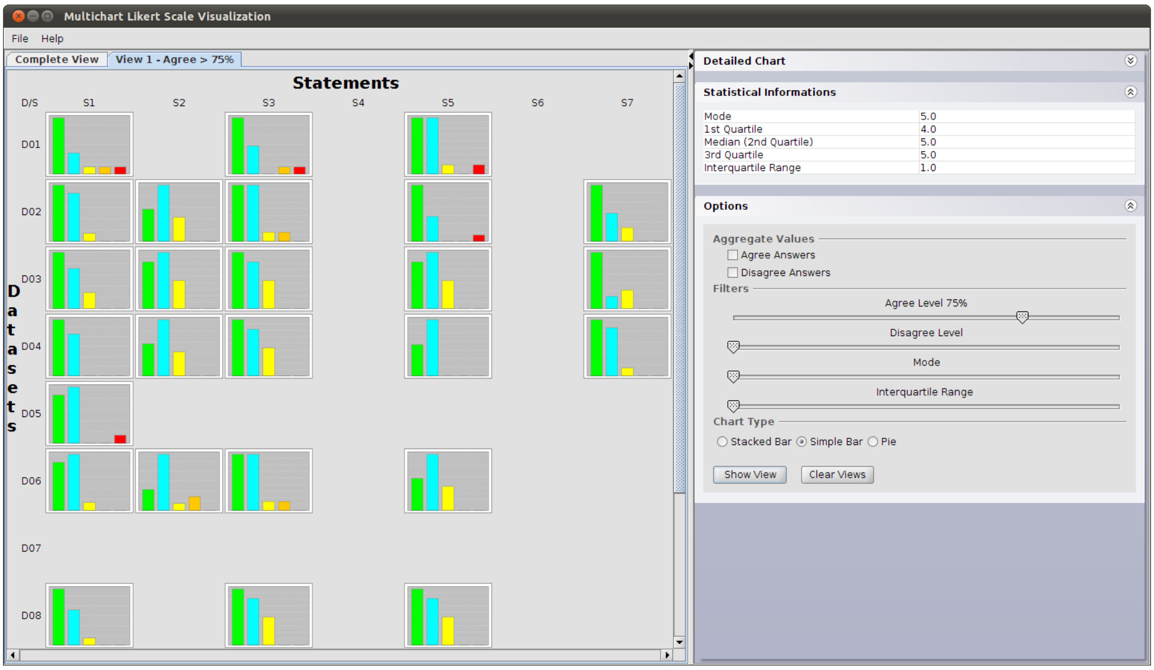

Figure 8: MCLikertVis - Filtered View.

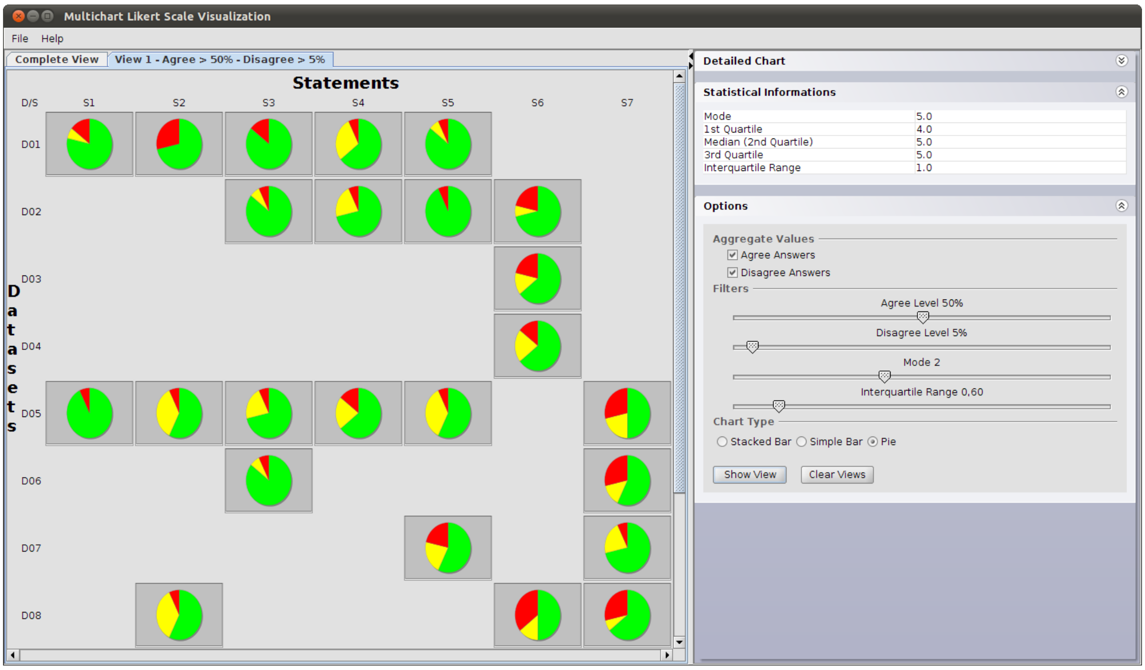

Figure 9: MCLikertVis - Pie chart and filtered view.

ading and parsing a data file generated by some tabulation tool (currently MCLikertVis reads well-formed XML files). With the data loaded, the tool performs the required calculation for plotting the data as well as basic statistics. The matrix of charts and the details area are displayed, and the tool is ready for user interaction. Figure 10 illustrates this process.

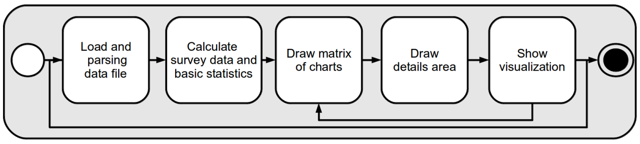

Figure 10: MCLikertVis Pipeline.

The parsing is performed on a proposed XML model represented in Figure 11. The *Survey* node consists of a set of *dataset* nodes, which represent the items that are being evaluated and a set of *Statement* nodes that are used in the evaluation procedure. The *answer* nodes are formed by the association between *dataset* and *statement*, storing the value from 1 to 5 (from Strongly Disagree to Strongly Agree).

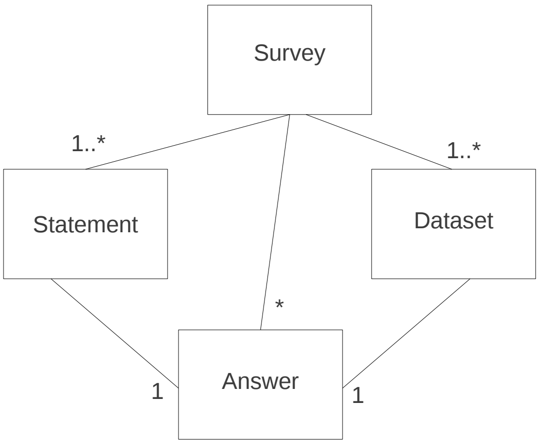

Figure 11: Abstract survey model.

4. CASE STUDY

In order to describe (and obtain a first evaluation of) the use of MCLikertVis, we performed a case study, where we evaluated a set of 11 graph layouts using a list of 7 affirmative statements and a 5-point Likert scale.

The case study is based on the subsequent work done on the MagnetViz technique that was originally presented by Spritzer and Freitas [12]. MagnetViz is a graph visualization technique that allows users to interactively manipulate force-directed graph layouts to make them tailored for their specific applications. This is done by means of virtual magnets, which can be placed anywhere on the scene and set to attract nodes and edges based on graph topology, the content of their attributes, or even on what other magnets attract or not. Figure 12 shows a graph visualization where some subgraphs are separated by means of virtual magnets.

In the original paper [12], MagnetViz's evaluation consisted only of a very brief assessment of the technique against a task taxonomy for graph visualizations [9]. Since then,

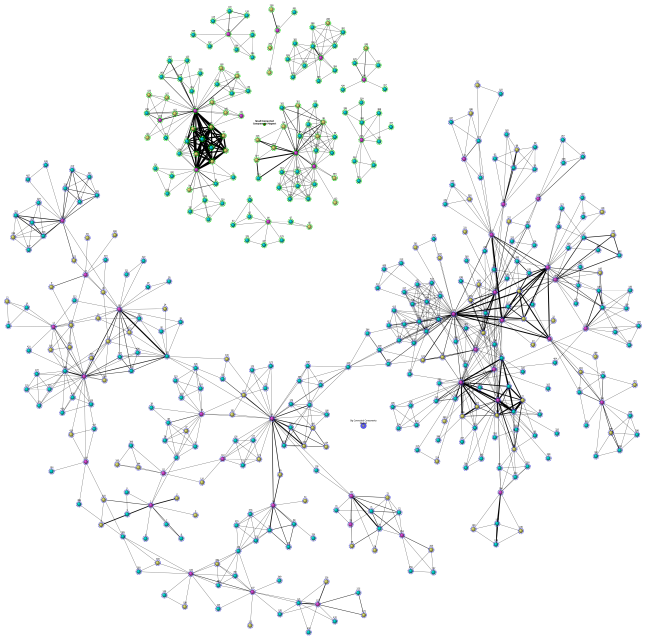

Figure 12: Graph visualization using MagnetViz.

MagnetViz underwent several changes and improvements, and was also subjected to a thorough validation process [11]. This process included an evaluation of aesthetic quality of the layouts generated with the technique, which was done by means of a user study.

The user study for evaluating MagnetViz's layout quality consisted of showing subjects an original force-directed layout and 10 modified versions made with MagnetViz. For each of the modified versions, users were given descriptions of what the image intended to portray. Subjects were asked to visually inspect each layout and rate them on a 5-point Likert scale according to their level of agreement to 7 affirmative statements.

To analyze the results, the mode and interquartile range of subject answers to each statement were computed for every layout, and bar charts were used to visualize subject answer distribution. One bar chart was drawn for each layout, with each statement being represented by bars of different colors. Initially, the bar charts were drawn as in Figure 4, but then, to allow for easier reading, charts were made as in Figure 4, in which agreement and disagreement answers were aggregated (i.e. "Strongly Agree" and "Agree" became a single "Agree" category, with the same being done for "Strongly Disagree" and "Disagree").

While by no means a bad visualization, in the case of the MagnetViz layout quality user study, these bar charts had some disadvantages from both practical and analytical standpoints. The practical disadvantages come from all the space these charts take up on a page, especially since each layout required a different chart. A large number of separate charts will make any evaluation report very difficult to analize. The analytical disadvantage is that the charts aggregate answers by layout, facilitating statement-to-statement comparisons within a single layout, but making it more difficult to see how a statement in a certain layout fares in relation to the same statement on another layout. Another main analytical disadvantage is that these charts are static, being

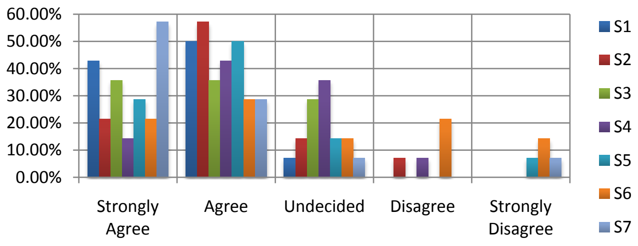

Figure 13: Bar charts used in the analysis of the answers to the MagnetViz layout quality user study.

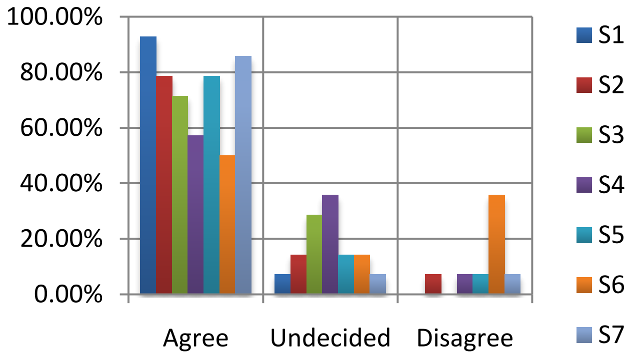

Figure 14: Aggregated bar charts for the same data used in Figure 4.

more useful to illustrate rather than to actually explore the data and make new discoveries. With MCLikertVis, all of these issues are dealt with.

MCLikertVis's approach of visualizing responses' distributions by means of an interactive matrix of charts is able to show simultaneously the answers to all statements for every layout, reducing page space requirements. The disposition of the answer distributions on a matrix also makes it easier to make comparisons, both regarding different statements on a same layout and the same statement on different layouts. This increase in ease of visual analysis is further supported by the exploratory tools. Aggregation cleans up the visualization, making certain aspects immediately clear, while filtering with the mode and interquartile range allows one to focus on the single layout-statements that had more or less variability in their answer distributions.

Figure 5 shows a stacked bar chart visualization generated by MCLikertVis representing the distribution of MagnetViz layout quality user study's answers. It can be seen from the outset for all charts that subjects were evenly distributed amongst all answer categories in statement 6. It can also be seen that the seventh layout (D07 in the picture) received in general more neutral and negative answers than the others. Overall, however, answers seem to have fallen mostly in the two agreement categories. This is made clearer in Figure 15, in which aggregation was turned on for both "Agree" and "Disagree" answers. Significant negative answers can be seen only in all statements concerning the seventh layout and in the sixth statement regarding each of the layouts. By turning off aggregation and filtering in this chart to find which statement answers had interquartile range higher than 2 (and thus a high variability), we end up with Figure 16. In this Figure it can be seen that regarding the sixth statement, subjects were almost evenly divided among positive and negative answers, with only a few marking the neutral answer. Still, a significant amount of people who marked "Strongly Disagree" only appeared in layout 07 (D07, statement 6). It can also be seen that in three layouts statement 7 received significant disagreement or neutral answers, despite leaning a bit more towards agreement. One can observe as well that while most subjects agreed with Layout 01 (D01, statement 2), a non-negligible number of people disagreed with it.

From what we have observed, MCLikertVis has fulfilled the needs of exploring the range of different answers.

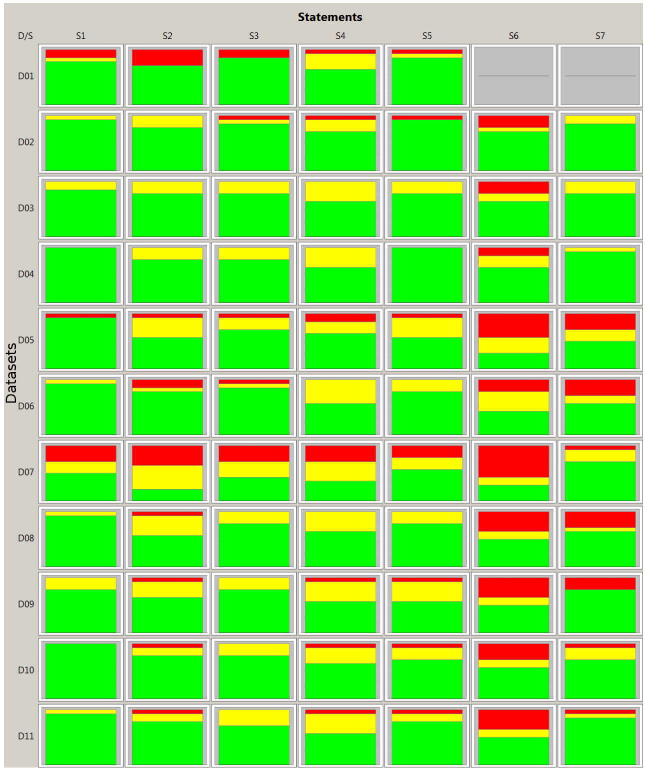

Figure 15: Answer aggregation in MCLikertVis on the MagnetViz dataset.

5. CONCLUSION

Likert scales are often adopted by psychologists, social researchers and also in any qualitative evaluation procedure. In general, Likert scales are easy to build and analyze, but usually that work is made with the help of traditional spreadsheets or data analysis tools, without the benefits of interactive visualization. Indeed, potential users of information visualization techniques often work with traditional statistical packages, and few know the benefits of interactive visualization.

An information visualization technique is meant to support the analysis and comprehension of (often large and/or complex) data sets in order to reveal or enhance features, patterns, clusters and trends, not always visible even when using a graphical representation. A practical example is the visualization of data obtained from user interface evaluation

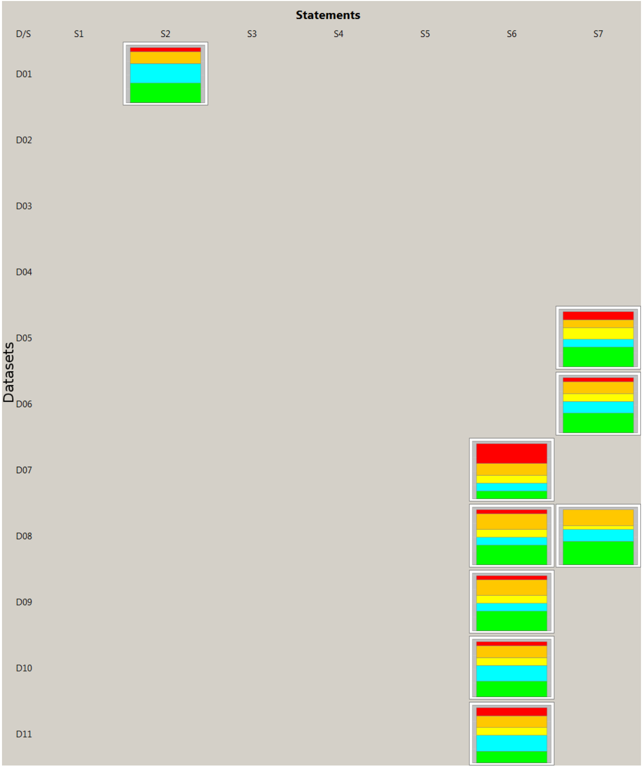

Figure 16: Filtering the MagnetViz dataset with the interquartile range.

methods, from which one would like to quickly identify potential usability problems and successful design choices.

In this paper we have presented an interactive visualization technique to help the analysis of Likert scale data. The main advantages of our tool is the possibility of comparing (using interactive features) the outcomes of the application of multiple instances of the same questionnaire, i.e., the same set of Likert items applied to different elements under evaluation. This work has also practical implications: to the best of our knowledge, there is no specific tool for the analysis and representation of Likert scale data. We have shown how to use MCLikertVis by presenting a case study where we analyzed the results of a user study on the layout quality of MagnetViz, a graph visualization tool [11].

MCLikertVis is particularly interesting for the analysis of data resulting from the evaluation of information visualization techniques because, although we agree with [8] and [5] concerning to the importance of quantitative evaluation methods, the evaluation process of such visualization techniques should follow rigorous human-computer interaction empirical methods with more culturally inspired approaches (see a discussion in [6] regarding this).

Future work include providing other kinds of charts as well as reordering lines and columns in the matrix of charts, like ScatterDice[7] does, so the user can put side by side two designs (lines) or two related Likert items (columns) and observe possible correlations. Finally, although we have successfully experimented MCLikertVis to analyze data from a qualitative evaluation procedure performed by Spritzer and Freitas [11], a thorough evaluation of MCLikertVis is our next step.

Acknowledgments

We thank CNPq - Brazilian National Research Funding Agency - for financing the work of C.M.D.S. Freitas and A. Spritzer.

6. REFERENCES

[1] Supersurvey. Online `http://www.supersurvey.com/`, may 2011.

[2] Surveygizmo. Online `http://www.surveygizmo.com/`, may 2011.

[3] Tableau software. Online `http://www.tableausoftware.com`, may 2011.

[4] Wammi. Online `http://www.wammi.com/`, may 2011.

[5] S. Carpendale. Information visualization. chapter Evaluating Information Visualizations, pages 19–45. Springer-Verlag, Berlin, Heidelberg, 2008.

[6] G. Cockton. Getting there: six meta-principles and interaction design. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pages 2223–2232, New York, NY, USA, 2009. ACM.

[7] N. Elmqvist, P. Dragicevic, and J.-D. Fekete. Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis 2008)*, 14(6):1141–1148, 2008.

[8] Y.-a. Kang, C. Gorg, and J. Stasko. Pragmatic challenges in the evaluation of interactive visualization systems. In *Proceedings of BELIV 2010*, BELIV'10, 2010.

[9] B. Lee, C. Plaisant, C. S. Parr, J.-D. Fekete, and N. Henry. Task taxonomy for graph visualization. In *BELIV '06: Proceedings of the 2006 AVI workshop on BEyond time and errors*, pages 1–5, New York, NY, USA, 2006. ACM.

[10] R. Likert. A technique for the measurement of attitudes. *Archives of Psychology*, 22 140:55–, 1932.

[11] A. Spritzer and C. D. S. Freitas. Design and evaluation of magnetviz - a graph visualization tool. *IEEE Transactions on Visualization and Computer Graphics*, Preprint.

[12] A. S. Spritzer and C. M. D. S. Freitas. A physics-based approach for interactive manipulation of graph visualizations. In *AVI '08: Proceedings of the working conference on Advanced visual interfaces*, pages 271–278, New York, NY, USA, 2008. ACM.