Are There Benefits in Seeing Double? A Study of Collaborative Information Visualization

Gloria Mark, Keri Carpenter, Alfred Kobsa
School of Information and Computer Science
University of California, Irvine, CA 94697
+1 949 824 7403
{gmark,keri,kobsa}@ics.uci.edu

ABSTRACT
We conducted an empirical study to better understand collaborative information visualization. We found that a system that offered fewer options for visualizations yielded more correct responses faster. Groups were more accurate but slower in solving problems than individuals. We identified different stages in visual discovery and found that collaboration benefits are from validating results and not from planning and system use. Tools to help translate and confirm the visualization would be of great benefit.

Keywords
Collaborative information visualization, CSCW, user study

INTRODUCTION
Recently there has been a surge of interest in the field of visual data mining and discovery, including collaborative information visualization (e.g. CVD and Cave6D [1]). Yet there is still a scarcity of user studies. The purpose of our paper is to investigate how groups make decisions about data using different information visualization systems, to help inform the design of such systems.

VISUALIZATION SYSTEMS USED
Subjects in our experiment used two different readily-available commercial systems for multivariate data visualization: Spotfire ((**SP**); formerly IVEE [1]) and InfoZoom ((**IZ**); formerly Focus [3]). SP offers users traditional visualizations such as scatterplots, bar charts, pie charts, graphs, parallel coordinates, and trellises. IZ presents information in table or distribution oriented formats. Users can sort data (possibly hierarchically), and "zoom" into information subspaces by double-clicking on attribute values or sets/ranges of values.

Fig. 1 shows portions of IZ's compressed mode, with data from a web-based dating service containing users' self-descriptions. The individuals are hierarchically sorted by their responses to the question "Do you participate in sports?" and, within each answer, by their weight in ascending order. The top line shows the weight distributions of these groups (there is a small third group of heavy people who did not answer this question). The bottom line shows the average weight per group: those who engage in sports are slightly heavier than those who do not. Figure 2 shows the same information in SP's bar chart visualization.

CHI 2003, April 5–10, 2003, Ft. Lauderdale, Florida, USA.
ACM 1-58113-630-7/03/0004.

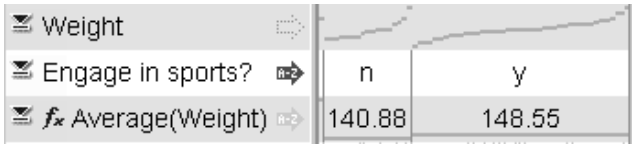

Figure 1: Infozoom's Compressed View

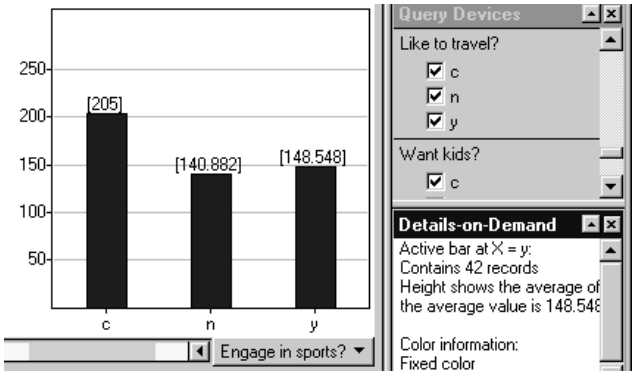

Figure 2: Spotfire's Bar Chart Visualization

THE EXPERIMENT
One hundred undergraduate students in Computer Science were randomly assigned in a 2 x 2 between-subjects design, to use either IZ or SP, and to work either alone or in pairs. To increase generalizability, groups worked either remotely using Netmeeting (N=20 pairs) or collocated using a SmartBoard (N=20 pairs). There were 20 subjects in the Alone condition. All subjects were videotaped and their screen activity was recorded using a screen capture utility.

Subjects used anonymized data from an online dating service to answer ten questions, such as: "Did males cheat more on their girlfriends than females on their boyfriends?" Each problem involved two to five variables. We measured the correctness of responses and time to completion, the latter from the videotapes. Two coders analyzed the group videotapes using ethnographic methods.

RESULTS

An ANOVA shows that subjects using IZ answered more questions correctly than with SP, $F(1,54)=4.1$, $p<.05$ (Table 1). As there was no difference between the remote and collocated groups, we combined their results. An ANOVA shows that being in a group yielded higher correct responses than working alone ($F(1,56)=3.4$, $p<.07$)**.** There was no significant interaction.

System / Interaction	InfoZoom	Spotfire	Total	N
Alone	6.8 (1.7)	6.7 (1.6)	6.8 (1.6)	20
Groups	8.2 (1.6)	7.0 (1.7)	7.6 (1.7)	40

Table 1: Mean and s.d. of correct responses

Table 2 gives the average time per problem, in seconds. An ANOVA showed that subjects using IZ were significantly faster solving the problems than subjects using SP ($F(1,58)=6.8$, $p<.01$). Remote and collocated showed no difference and when combined, an ANOVA showed groups were slower than individuals ($F(2,58)=5.6$, $p<.02$.

System / Interaction	InfoZoom	Spotfire	Total	N
Alone	114.6 (28.3)	124.1 (32.4)	119.6 (30.0)	20
Groups	130.1 (61.2)	185.0 (71.2)	157.6 (71.2)	40

Table 2. Avg. time and s.d. per problem (in seconds)

A model of collaborative information visualization

The video analysis revealed a very consistent sequential pattern for solving focused information visualization tasks. We explain our quantitative results by showing how different stages were affected by the conditions (Fig. 3).

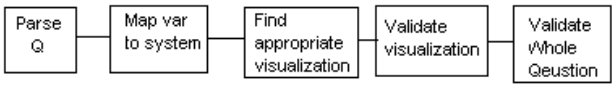

Figure 3. Stages of information visualization

In the first stage, subjects read the question and split it into distinct variables. In stage two, variables from the question were mapped to the visualization system. IZ users made far fewer errors in this step because all variable names are visible in a list on the left-hand side of the screen. SP requires more effort in this stage as relevant variables can appear in either or both of two screens – oftentimes relevant variables must be scrolled through on the top right section of the screen. Searching for variables in multiple views can explain why SP users were slower with more errors.

In stage three, users manipulated the system to find the right visualization for the variables. SP users took longer than IZ users in this stage as they had more options for visualizations. More choices led SP users to make more errors. With IZ, users had only one view of the data and they therefore had fewer problems manipulating the system.

In stage four, users validated the visualization. If there was a discrepancy, they discussed it. If the person with the cursor got lost or the display was incorrect, she often gave control of the cursor to the other partner. It is this stage where groups showed the most benefit over individuals: they were more correct yet slower due to confirming that their responses were correct. We observed no differences with SP and IZ use in this stage. Stages two through four were then repeated for each additional variable.

In the last stage, the users validated the entire answer. Verbal interaction also occurred at this stage, but little system use. Participants confirmed with their partner that the visualization (and thus their answer) satisfied every aspect of the question.

DISCUSSIONS AND CONCLUSIONS

Our data shows working with SP yielded more incorrect responses and slower times. While both systems allow the same discoveries, SP offers more options and solution paths than IZ, while IZ is more transparent by making more data and functionality easily visible and accessible.

Groups tended to yield more correct responses than individuals but were slower. Contrary to expectation, both IZ and SP groups tended not to negotiate a specific plan of action before beginning and interacted very little in choosing variables to visualize with the system. To our surprise, we found that the benefits for collaboration in information visualization occurred in the confirmation stage, and not in the earlier planning or system interaction stages. In almost all groups, one subject had cursor control. Most verbal group interaction occurred during the validation stages: for each variable, and at the end with the visualization for answering the entire question. Thus, the main advantage of groups over individuals appears to be in confirming the answer; it led them to more correct answers but took longer. A trend showed that groups did not perform better than individuals with a less transparent system, as with SP. The implications for HCI are that the increased accuracy yet longer time taken by two people may not pay from a cost/benefit perspective. Tools that help users translate and confirm the visualization with the problem variables would be of immense help.

REFERENCES

1. Ahlberg, C. and Wistrand, E. (1995). IVEE: An Information Visualization and Exploration Environment, *Proc. of InfoVis'95*, New York, NY, 66-73.
2. Lascara, C., Wheless, G., Cox, D., Patterson, R., Levy, Johnson, S. A. E. and Leigh, J. (1999). "TeleImmersive Virtual Environments for Collaborative Knowledge Discovery", *Advanced Simulation Technologies Conference*, San Diego, CA, 1999
3. Spenke, M., Beilken, C. and Berlage, T. (1996). "The Interactive Table for Product Comparison and Selection", *UIST '96 Ninth Annual Symposium on User Interface Software and Technology*, Seattle, 41-50.