IADIS International Journal on Computer Science and Information Systems
Vol. 12, No. 2, pp. 17-31
ISSN: 1646-3692

RADIAL VS. RECTANGULAR: EVALUATING VISUALIZATION LAYOUT IMPACT ON USER TASK PERFORMANCE OF HIERARCHICAL DATA

Sujay Muramalla[1], Ragaad AlTarawneh[2], Shah Rukh Humayoun[3], Ricarda Moses[1], Sven Panis[4] and Achim Ebert[1]
[1]Computer Graphics and HCI Group, University of Kaiserslautern, Germany
[2]Information Technology Department, Mutah University, Jordan
[3]Computer Graphics and HCI Group, University of Kaiserslautern, Germany
4Faculty of Social Sciences, Experimental Psychology Unit, University of Kaiserslautern, Germany

ABSTRACT

Space-filling techniques have been used in the information visualization field as an alternative to the conventional node-link layouts for intuitively showing large hierarchies in less space. Different space filling layouts have been designed, developed and evaluated; however, much less effort have been made to look into how layout can impact user task performance on hierarchical data structures. In this paper, we focus on the impact of layout on user task performance by conducting evaluation studies for two common space-filling layout structures, the Sunburst (radial) layout and the Icicle (rectangular) layout. In our studies, users performed eight search-based tasks on files and directories in the resulting visualizations, first in a controlled environment and subsequently in an online environment. We focused on deriving user performance metrics with regard to effectiveness, efficiency, and user acceptance. Results demonstrate a mixed view of task performance and preference with both layouts, e.g., users performed better with the Icicle layout while they preferred the Sunburst layout for visual aesthetics. We further analyzed the impact of layout on the performance dynamics in terms of response times and accuracy using event history analysis (EHA) in the control study setting. The EHA results revealed clear differences in response tendencies even though no differences existed in mean response times for most of the tasks. It also clearly showed that participants performed more efficiently with the directory comparison tasks than the file comparison tasks. Overall, through these studies we were able to derive causal relationships between the layout and the user's task performance while interacting with hierarchical data structures.

KEYWORDS

Usability Evaluation, Space-Filling Techniques, Information Visualization, Hierarchical Data Structures, Sunburst Layout, Icicle Layout

1. INTRODUCTION

In the case of visualizing hierarchical data, one of the main concerns is to represent large tree data structures in a limited visual display area. Therefore, many techniques have been proposed in the state-of-the-art literature to visualize this kind of data using different representation forms such as tabular, textual, node-link diagrams, or space-filling techniques [11, 12, 24, 25]. The differences in these tree visualization techniques lead to perceiving them differently by different viewers due to the variations in their cognitive abilities. For example, finding the root of the tree is considered as the initial task to understand this representation structure. However, tree visual representations make this simple task quite challenging for the non-expert viewers. For example, the root is assumed to be in the center of the radial tree layout but in the case of large tree sizes, high number of edges between the tree levels makes it difficult to find the root for many non-expert users.

Most of the tree drawing algorithms utilize edges between the nodes to show the parent-child relation. This can lead to the problem of overlapping for large tree sizes. To deal with such a problem, different space-filling visualization techniques (e.g., the Tree-map layout, the Sunburst layout, or the Icicle layout [12, 25]) have been proposed to produce compact representations for large tree sizes. Such techniques were designed using geometrical features and solid areas to depict hierarchical structures. More specifically, these space-filling visualization techniques resulted in three different space-filling methods: the rectangular (TreeMap), the Icicle, and the circular (Sunburst). Nowadays, these techniques are especially popular for visualizing large trees to represent software systems, social networks of large communities, etc. Therefore, it is critical to understand how users perceive the different space-filling approaches in regard to accuracy, efficiency, and acceptance in order to utilize them efficiently and effectively.

Keeping this goal in mind, we performed user evaluation studies on the two selected space-filling visualization layouts (i.e., the Sunburst lay and the Icicle layout) to assess user task performance while interacting with a hierarchical data structure in terms of accuracy and efficiency. Users performed eight tasks on the visualizations representing the hierarchical structure of a software system. The first user study was executed in a controlled laboratory setting while the second user study was executed in an online setting. In the controlled laboratory setting, we also collected user acceptance feedback using closed-ended and open-ended questionnaire forms. In the online setting, we collected user rating regarding their preferred layout between the two visualization layouts.

Results from the both studies demonstrate a mixed view of task performance and preference with both layouts, e.g., users performed better with the Icicle layout while they preferred the Sunburst layout for visual aesthetics. We further analyzed the impact of layout on the performance dynamics in terms of response times and accuracy using event history analysis (EHA) in the control study setting. The EHA results revealed clear differences in response tendencies even though no differences existed in mean response times for most of the tasks. It also clearly showed that participants performed more efficiently with the directory comparison tasks than the file comparison tasks. Overall, through these studies we were able to derive causal relationships between the layout and the user’s task performance while interacting with hierarchical data structures. The results of both studies would help the researchers and practitioners in selecting one of the options of space-filling visualization techniques that would suit their specific needs.

The remainder of the paper is structured as follows: In Section 2, we briefly provide the related work. In Section 3, we highlight the software structure and explain the two research experiment settings. In Section 4, first we provide results of the both conducted studies in the terms of accuracy, efficiency, and user acceptance, then we explain the event history analysis results, followed by general findings. Finally, we conclude the paper and highlight the future directions in Section 5.

2. RELATED WORK

Hierarchical information structures are one of the commonly used information structures [26]. Information structures such as family trees, building plans, computer directory structures, museum information, art galleries, libraries, or cataloging are normally represented as hierarchical data structures [12, 24]. The need for visualization for depicting hierarchical information was motivated by early computing challenges of visualization of large scale data structures stored in hard drives [12]. Johnson and Schneiderman [12] showed a TreeMap visualization layout in their study that maps hierarchical information structure to a 2-D space filling rectangular display while utilizing 100 percent of the available area. They also highlighted the importance of TreeMap over traditional tree drawing techniques such as lists, outlines, and graph trees, which were suitable for small hierarchies but failed in depicting large scale hierarchies [12]. Turo and Johnson [26] presented TreeMap in two different ways, i.e., the Top-Down and the Slice and Dice. While the top-down TreeMap method preserved traditional tree diagram convention fostering comparative analysis, it failed in depicting large scale hierarchical data [26]. The slice and dice TreeMap was beneficial for efficiently displaying large scale hierarchies; however, it recursively partitioned in both the dimensions [26]. They further performed an empirical evaluation of the TreeMap algorithm by comparing it with UNIX for carrying out directory browsing tasks [24, 26]. Results of the evaluation study favored UNIX; however, the authors attributed the difference due to expert UNIX participants. Stasko et al. [24] later conducted an evaluation study for space filling visualization tools for depicting hierarchical data structures by examining two methods of space filling, i.e., TreeMap (rectangular) and Sunburst (radial). In their evaluation study, results favored the Sunburst layout in the case of performance. While González-Torres et al. [8, 9] investigated how users perceive various visualization techniques to depict large scale hierarchical information structures by performing various tasks applicable to the evolution of large scale software files and directory structures. Recently, AlTarawneh and Humayoun [4] provided enhancements for the Sunburst layout in order to show the overall software system structure and the inside details in a compact visual form. They also provided the initial findings of their preliminary user evaluation study. Finally, AlTarawneh et al. [5] proposed some optimization solutions and interaction options for the Sunburst visualization layout to make it appropriate for mobile devices. They performed a user evaluation study with 15 participants to get their feedback regarding the influence of adding the colorization styles to convey the hierarchy clearly.

3. THE EVALUATION STUDIES

It is essential that the visual designs agree with the cognitive framework supporting UI dimensions for intuitive interaction [1, 8]. Also, it is important to understand that a cognitive study is necessary to reduce the gaps between the UI and the system [1]. Therefore, support for cognitive based design is inevitable [8]. This helps in more effortless recall, search, and browse functions of UI dimensions as it is easier to learn and is more satisfying [6, 13, 18]. In our studies, we aimed at studying the impact of visual layout on user task performance in terms of accuracy, efficiency and acceptance. Visual layouts are important elements of visual hierarchy and play a crucial role in task selection and execution during search and scan processes of interaction. In the following subsections, we explain our experimental design and settings.

3.1 The Two Experiments Settings

We used the software structure hierarchy of the RAVON [8] robot in the visualizations central in our studies. The RAVON software system contains four main packages (i.e., libraries, projects, tools, and others), where each package contains a group of sub-packages except the "other" package as it consists of less than 0.6% of the RAVON software system. The "libraries" package is the largest one (approximately 55.7%) in the RAVON software system, followed by the "projects" package (approximately 33.1%), and the "tools" package (approximately 10.7%). In total, RAVON has 3152 software files and are characterized by three extensions (i.e., ".cpp", ".h", and ".hpp"). Each one of these files represents a leaf node in the hierarchy. Further, the RAVON software files are divided by their levels with a maximum level is of six degrees.

Figure 1 shows the resulting Sunburst and Icicle visualizations of the RAVON software structure hierarchy, which were developed using the D3[1] library. The color scheme in both visualizations were inspired from the work of AlTarawneh et al. in [4, 5]. The visualizations were positioned in the central portion of the screen, whereas the activity questionnaire for users to solve tasks appears in the frame on the left side of the screen. Also, the labelling scheme providing contextual information to the user appears at the top of the screen.

The experiment was a between-subject design, where half of the randomly selected participants performed the tasks with the help of the Sunburst visualization and the other half with the Icicle visualization. We conducted the experiment based on two evaluation settings. The first was a controlled laboratory setting where the moderators were present to note down the data and to answer the questions in case of any difficulties faced by the participants. Figure 2 shows the controlled experiment setup in the lab environment. The second was an online study setting conducted by hosting the website online via an FTP server. The webpage links for both Sunburst and Icicle were shared through emails, blogs, and social media. In this evaluation setting, the experiment was performed without any moderators.

[1] D3 Data-Driven Documents: http://d3js.org/

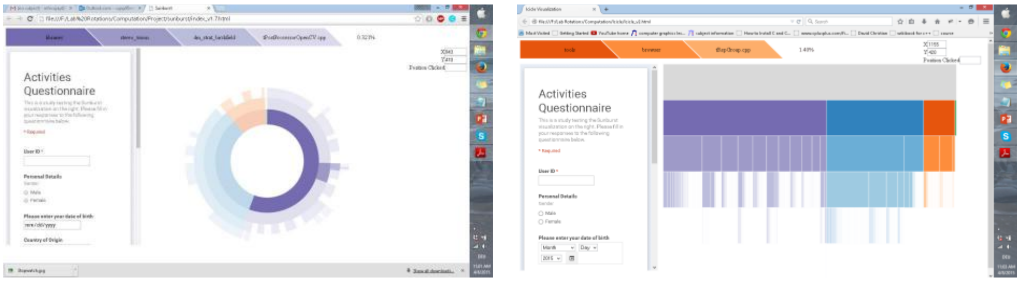

Figure 1. Sunburst (left) and Icicle (right) with activity questionnaires for the participants on the left frame in both visualization tools

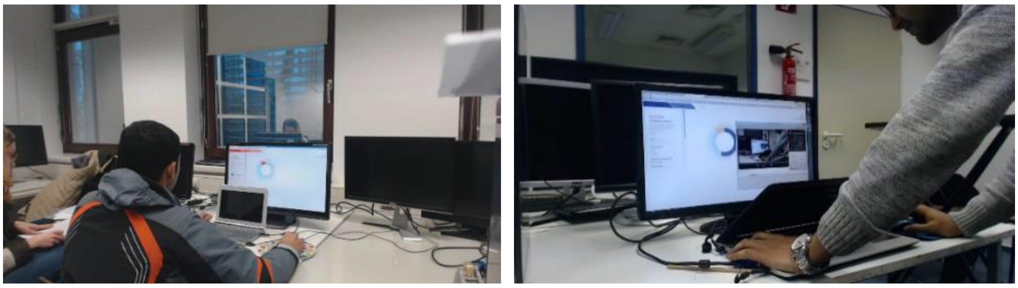

Figure 2. Participants performing in the controlled experiment

3.2 Tasks and Hypotheses

A set of 8 activity tasks were designed to extract user performance measurements from both visualization layouts. In both evaluation settings, participants were asked to complete the following 8 tasks:

1. Identify the largest directory in the hierarchy
2. Identify the largest file in the libraries directory
3. Identify the second largest directory in the hierarchy
4. Identify the second largest file in the projects directory
5. Identify the deepest subdirectory of the directory "/libraries/stereo_vision"
6. Compare the directories "/projects/ravon/control/" and "/projects/ravon/navigator/". Which one is the larger one?
7. Compare the directories "/projects/ravon/obstacle_detection/" and "/projects/ravon/navigator/". Which one has more files?
8. Which directory has the least amount of levels?

At the end of the lab study, participants were asked to give their feedback regarding their acceptance level using a closed-ended questionnaire form using a Likert-scale and an open-ended questionnaire form. In the case of the online study, participants were asked at the end of the test to rate their preferred visualization layout.

In the study by Stasko et al. [25], the authors found a preference for and a higher subjective opinion of the Sunburst visualization layout in comparison to the TreeMap visualization layout. In terms of perception of symmetry, we believe that the Sunburst with a radial layout is more symmetric and hence visually aesthetic. However, the rectangular layout of the Icicle layout supports the top-down viewing hierarchy; therefore, we hypothesize that the top-down viewing preference lets the participants execute the task more accurately. Further, since the Icicle layout follows the conventional tree structures; therefore, we assume that the visual hierarchy and the linear navigation of the rectangular display layout (i.e., Icicle in our case) will support participants in executing tasks more quickly. Based on these assumptions, we formulated the following set of hypotheses:

1. **Hypothesis 1:**
   Accuracy (Icicle Layout) >= Accuracy (Sunburst Layout)
2. **Hypothesis 2:**
   Efficiency (Icicle Layout) >= Efficiency (Sunburst Layout)
3. **Hypothesis 3:**
   User Acceptance (Sunburst Layout) >= User Acceptance (Icicle Layout)

3.3 Participants

Participants were divided into two separate groups with each group performing tasks on one visualization (either Sunburst or Icicle). Four participants performed the pilot study. In the case of the lab experiment, 36 participants (10 females and 26 males) took part with a mean age of 27.09 years (range 24 – 32 years). Most of the participants were students with different cultural and educational backgrounds. All the participants who performed the experiment were familiar with the notion of files and directories, but interaction with interactive visualization layouts like Sunburst and Icicle were new to them. In the online study, 59 participants (21 females and 38 males) took part. We randomly assigned the participants to one of the underlying visualizations to avoid any influence of demographic differences. Here, the participants' mean age was 27.96 years (range: 18 years – 51 years), with Sunburst having a mean age of 27.54 years (range: 18 years – 51 years, SD: 6.74 years) and Icicle a mean age of 28.32 years (range: 21 years – 38 years, SD: 4.50 years).

4. RESULTS AND DISCUSSION

In this section, we describe the results of the controlled lab study and the online evaluation study for the Sunburst layout (SB) and the Icicle layout (IC). First, we focus on the comparison of these two layouts in terms of accuracy, efficiency, and user acceptance. Then we explain the Event History Analysis (EHA) results on the response times and accuracy data to study the impact of layout on the temporal dynamics of user performance for each task. Finally, we discuss our overall findings.

4.1 Results of Accuracy, Efficiency, and User Acceptance

The average accuracy obtained from the controlled, online, and combined cases are represented in Table 1 and Figure 3. Data for accuracy results was tested for normality using the Shapiro-Wilk test statistic (SB: w=0.88, IC: w=0.97) for the average task scores obtained from each task when combined for both control as well as the online study. The results were then tested using the t-test for the reliable differences. The results show that users tend to be more accurate with the Icicle layout (IC) than the Sunburst (SB) layout. In the controlled study, the difference was closely significant using the t-test (IC=87%, SB=70%, and p=0.07). In the online study, the difference was not significant (IC=79%, SB=75%, and p=0.74). Additionally, the difference was also not significant when both cases were combined for average accuracies (IC=83%, SB=73%, and p=0.22). Overall, although, there were no significant differences, there seemed to be a higher accuracy trend using the Icicle layout than the Sunburst layout as shown in Figure 3.

Table 1. Participants' accuracy results in both studies as well as the combined results.

Layout	Tasks ID								
	1	2	3	4	5	6	7	8	Average
	Controlled Lab Study								
Sunburst	0.88	0.50	1.00	0.16	0.77	0.94	1.00	0.33	**0.70**
Icicle	0.94	0.61	0.94	0.77	0.94	1.00	1.00	0.77	**0.87**
	Online Study								
Sunburst	0.88	0.60	0.92	0.24	0.76	0.88	1.00	0.80	**0.75**
Icicle	0.74	0.77	0.77	0.70	0.85	0.96	0.74	0.77	**0.79**
	Average Results of Both Studies								
Sunburst	0.88	0.55	0.96	0.20	0.77	0.91	1.00	0.57	**0.73**
Icicle	0.84	0.69	0.86	0.74	0.90	0.98	0.87	0.77	**0.83**

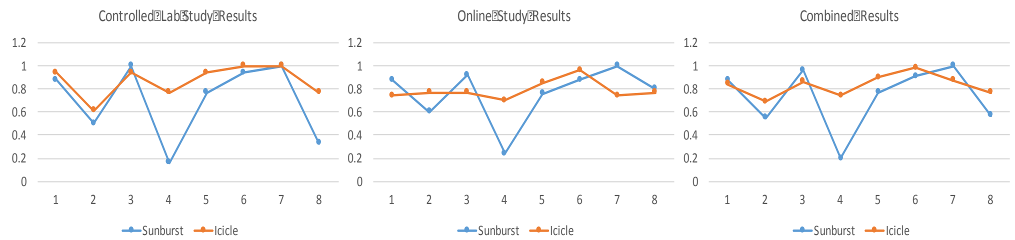

Figure 3. Participants' accuracy results in both studies as well as the combined results

The average efficiency results from the controlled, online and combined cases were measured as the average time taken to correctly completing the tasks and are shown in Table 2 and Figure 4. Because of varying complexity of each task, the average time of completion differs between the tasks. Data was tested for normality using the Shapiro-Wilk test statistic for the controlled lab study (SB: w=0.94, IC: w=0.87) and was found to be normally distributed. The average results were then tested using the t-test for the reliable differences. The results trend show that users tend to be quicker with the Icicle layout (IC) than the Sunburst (SB) layout. All the differences in average scores were closely significant with the t-test. In the controlled study, participants on average were quicker to complete the tasks with the Icicle layout (IC=43.67s, SB=55.17s, and p=0.11). In the online study, participants on average were quicker to complete all the tasks with the Icicle layout (IC=48.81s, SB=123.56s, and p=0.12). Furthermore, by combining both the results, participants on average were quicker to complete all the tasks with the Icicle Layout (IC=46.24s, SB=89.37s, and p=0.10). Although, there were no real significant differences; however, there seemed to be a higher efficiency trend using the Icicle layout than the Sunburst layout as shown in Figure 4.

Table 2. Participants' efficiency in seconds in both studies as well as the combined result

Layout	Tasks ID								
	1	2	3	4	5	6	7	8	Average
	Controlled Lab Study								
Sunburst	32.56	85.33	18.94	109	67.57	59.88	43.56	24.5	**55.17**
Icicle	19.06	96	23.11	62	58.23	34.06	37.56	19.36	**43.67**
	Online Study								
Sunburst	54.55	252	23.48	420	88.42	60	48	42	**123.56**
Icicle	39	88.57	22.86	75.79	57.39	30	54	22.86	**48.81**
	Average Results of Both Studies								
Sunburst	43.56	168.67	21.20	264.50	78	59.94	45.78	33.25	**89.37**
Icicle	29.03	92.29	22.99	68.90	57.81	32.03	45.78	21.11	**46.24**

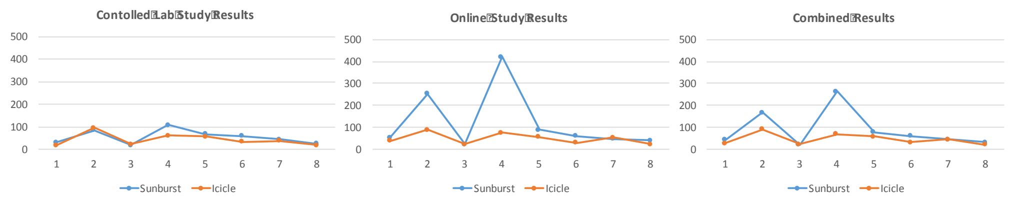

Figure 4. Participants' efficiency results in seconds in both studies as well as the combined results

In the controlled study, we asked 13 questions in the closed-ended questionnaire form (see Table 3) in order to know the participants' subjective preference for various features of the underlying visualizations. The table shows the frequencies with which each Likert score was selected by the participants. None of the differences were significant. However, there was a marginal difference in average satisfaction rating in the controlled study in favor the Sunburst layout.

In the online study, only one question was used to study the participants' acceptance. The difference in frequency means was found to be highly significant with a higher preference for the Sunburst layout, with 3.8 vs 3.1 and p-value of 0.0394. Table 4 shows the results of the online study.

In both studies, the number of positive responses on the Likert scale (count of 4s and 5s) were higher with the Sunburst layout compared to the Icicle layout. Additionally, this was also evident from the open-ended feedback from the participants, as they provided more positive feedback for the Sunburst layout compared to the Icicle layout.

Table 3. Participants' acceptance feedback in closed-ended questionnaire form in the controlled lab study

Questions	SUNBURST							ICICLE						
	1	2	3	4	5	Don't Know	Median	1	2	3	4	5	Don't Know	Median
1. I was able to compare the sizes of files using the labeling system.	0	3	3	7	5	0	4	0	2	4	5	7	0	4
2. I was able to figure out the largest file using the labeling system.	1	1	4	3	9	0	4.5	0	1	2	8	6	0	4
3. I was able to figure out the largest file using the size of the node.	1	2	4	5	6	0	4	0	1	5	5	7	0	4
4. I was able to figure out which subdirectories were inside another directory using Sunburst.	0	2	3	6	7	0	4	0	1	1	12	4	0	4
5. The coloring scheme was helpful while searching a particular file	2	0	3	5	8	0	4	2	4	1	3	6	2	4
6. The labeling system was helpful while searching for a particular file	0	4	4	6	4	0	4	0	2	5	5	6	0	4
7. The coloring scheme was helpful while searching for a particular directory	0	2	4	5	7	0	4	2	3	4	4	5	0	3.5
8. The labeling system was helpful while searching for a particular directory.	0	2	2	10	4	0	4	0	2	2	7	7	0	4
9. After completing the tasks, I now know how to use Sunburst well.	0	2	2	8	6	0	4	0	2	5	6	5	0	4
10. There are definitely times that I would like to use Sunburst.	2	1	4	8	3	0	4	0	0	9	7	2	0	3.5
11. I found the tool to be confusing to use.	1	10	3	1	3	0	2	3	5	5	5	0	0	3
12. Overall, I liked the tool.	0	2	3	11	2	0	4	0	2	5	9	2	0	4
13. I would like to use the tool to visualize my own hierarchical data.	0	4	4	8	2	0	4	0	8	2	7	1	0	3

Table 4. Participants' acceptance feedback in closed-ended questionnaire form in the online study

Visualization	How was the interaction with the tool?							
	1	2	3	4	5	Don't Know	Median	Mean
Sunburst	1	2	6	12	7	3	4	3.79
Icicle	4	7	8	6	6	0	3	3.10

4.2 Results of Event History Analysis

We performed the Event History Analysis (EHA) on the response times (RT) and accuracy data in order to study the impact of layout on the temporal dynamics of user performance for each task. Event history analysis is the standard distributional method for analyzing time-to-event data in many scientific disciplines. It is also known as survival, hazard, duration, failure-time, or transition analysis [2, 3, 10, 16, 17, 19, 20, 22, 23].

It is assumed that for each time point from the onset of the task, for each trial by each participant, there is a tendency or risk for the event (here, a response) to occur. The function relating this likelihood of response occurrence with the passage of time is known as the continuous-time hazard function of the response occurrence [7, 15].

In this study, we applied discrete-time (descriptive and inferential) methods [3, 20, 23, 27]. We divided the first 80 seconds beginning from task-onset into 8 time bins of 10 seconds each (i.e., (0,10], (10,20], (20,30], ... , (70,80] indexed by t = 1, 2, 3, ..., 8), and then estimated the discrete-time hazard function of response occurrence, h(t):

$$h(t) = P(T = t \mid T \geq t)$$

Here, $T \geq t$ denotes the event that the response did not occur before the start of time bin t. Thus, this conditional probability function gives each bin t the conditional probability of a response occurrence sometime during bin t, given that the response has not yet occurred in any previous time bin (t-1, t-2, ..., 1). The estimate is obtained by dividing the number of events by the risk set for each bin (see Table 5).

The survivor function, S(t), expresses the probability that the response has not yet occurred by the time bin t is completed. It is the joint probability that the response has not occurred in any of the bins prior to and including t:

$$S(t) = P(T > t) = [1 - h(t)] \cdot [1 - h(t\text{-}1)] \cdot [1 - h(t\text{-}2)] \cdot ... \cdot [1 - h(1)]$$

Finally, P(t) = P(T = t) indicates the unconditional probability that the response occurs in time bin t. Plotting P(t) over time t gives the (sub)probability mass function of response occurrence. The estimate of P(t) is obtained by dividing the number of events in each bin by the maximal risk set (i.e., 18; see Table 5). In Table 5 we illustrate the calculation of these functions for one condition (i.e.: Sunburst, Task 1).

Table 5. Life table for the condition of the Sunburst layout in Task 1 in the control study setting

T	#censored	#events	Risk Set	h(t)	S(t)	P(t)	Ca(t)
1	0	1	18	0.06	0.94	0.06	1.00
2	0	4	17	0.24	0.72	0.22	0.75
3	0	6	13	0.46	0.39	0.33	0.83
4	0	3	7	0.43	0.22	0.17	1.00
5	0	3	4	0.75	0.06	0.17	1.00
6	0	0	1	0.00	0.06	0.00	NA
1	0	1	18	0.06	0.94	0.06	1.00
2	0	4	17	0.24	0.72	0.22	0.75

In the study we have two event types of interest: correct and incorrect responses. We cannot simply estimate the hazard functions for correct and incorrect responses separately, because these two events cannot be assumed to be independent of each other [19]. Therefore, we take the conditional-processes approach by extending the h(t) analysis of response occurrence with the analysis of conditional accuracy [3]. By combining the hazard function with a conditional accuracy function, we can provide an unbiased, time-varying, and probabilistic description of the latency and accuracy of responses. First, we estimate h(t) of response occurrence regardless of response accuracy, to study whether and when responses occur. For each time bin t, the sample-based estimate of h(t) is obtained by dividing the total number of observed responses in bin t by the risk set for bin t. The risk set equals the total number of trials that are response-free in all time bins earlier than t and are thus still eligible to experience the response. Second, once we know probabilistically whether and when responses occur, we estimate the conditional accuracy function of observed responses, ca(t):

$$ca(t) = P(\text{response correct} \mid T = t)$$

Here, ca(t) provides the conditional probability that a response is correct given that it occurs sometime during bin t. The conditional accuracy, ca(t), is obtained by dividing the number of correct responses observed during bin t by the total number of observed responses in bin t. In conclusion, we took the following metrics into account for analyzing our time-to-event data:

1. Hazard function
2. Survival function
3. (Sub) probability mass function
4. Conditional accuracy

In Figures 5 and 6, we plot these sample-based functions for each task in the control study. For example, given that no response has occurred during the first 10 seconds in the Icicle layout of Task 1, there is a hazard probability of about 0.5 that the response is going to occur in bin (10, 20] (see Figure 5, upper left plot, bin endpoint 20); and given that a response occurs in this condition in bin (10, 20] it is estimated to be correct 100 percent of the time (see Figure 5, lower left plot, bin endpoint 20).

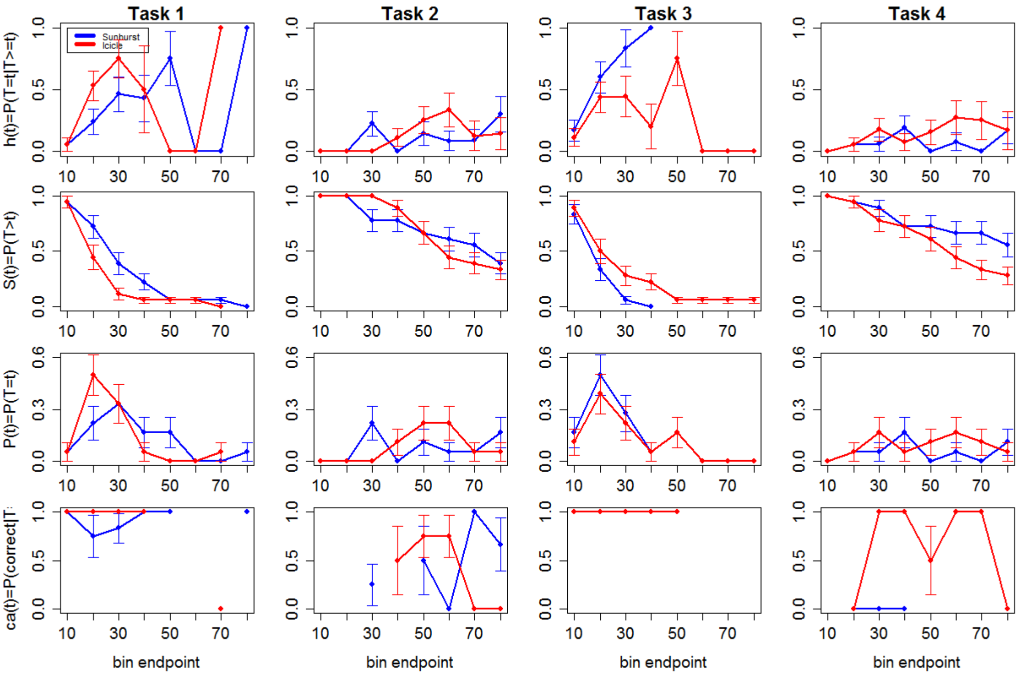

Figure 5. Sample-based descriptive statistics for Tasks 1 to 4 in the control study setting. Hazard estimates h(t), survival estimates S(t), (sub)probability mass estimates P(t), and conditional accuracy estimates ca(t) for the both layouts are presented in rows 1, 2, 3, and 4 respectively

More generally, the hazard and conditional accuracy functions show that for Tasks 1, 4, 5, 6, 7 and 8, there are more early responses for the Icicle layout compared to the Sunburst layout, and that when such an early response occurs, it tends to be more often correct for the Icicle layout compared to the Sunburst layout.

For Task 2 we observe a positive hazard for the Sunburst layout in bin (20, 30] but when a response occurs in bin (20, 30] it is estimated to be correct only 20 percent of the time. For later bins, hazard is higher for the Icicle layout and these later responses are more often correct for the Icicle layout compared to the Sunburst layout. For Task 3, hazard is higher for the Sunburst layout, but no differences in conditional accuracy are observed. In other words, with both layouts correct responses were given before 80 seconds after task onset. Interestingly, for Tasks 1, 4, 6, and 8 we see that early responses are less accurate for the Sunburst layout compared to the Icicle layout.

One of the advantages of using EHA is that it takes into account right-censored observations, in contrast to mean RT performance measures. For example, for Tasks 1, 2, 3, 5, and 7 we observed no differences in mean RT, but clear differences between the hazard functions of both layouts. We also noticed that for Tasks 1, 3, and 8 people were very fast, while in Tasks 2 and 4 people were very slow (regardless of layout). This indicates that tasks involving directory comparisons can be performed faster (refer to ca(t) plots for the tasks in Figures 5 and 6) than tasks involving file comparisons. The data for the online study setting was not analyzed using EHA because the time measurements were too unreliable.

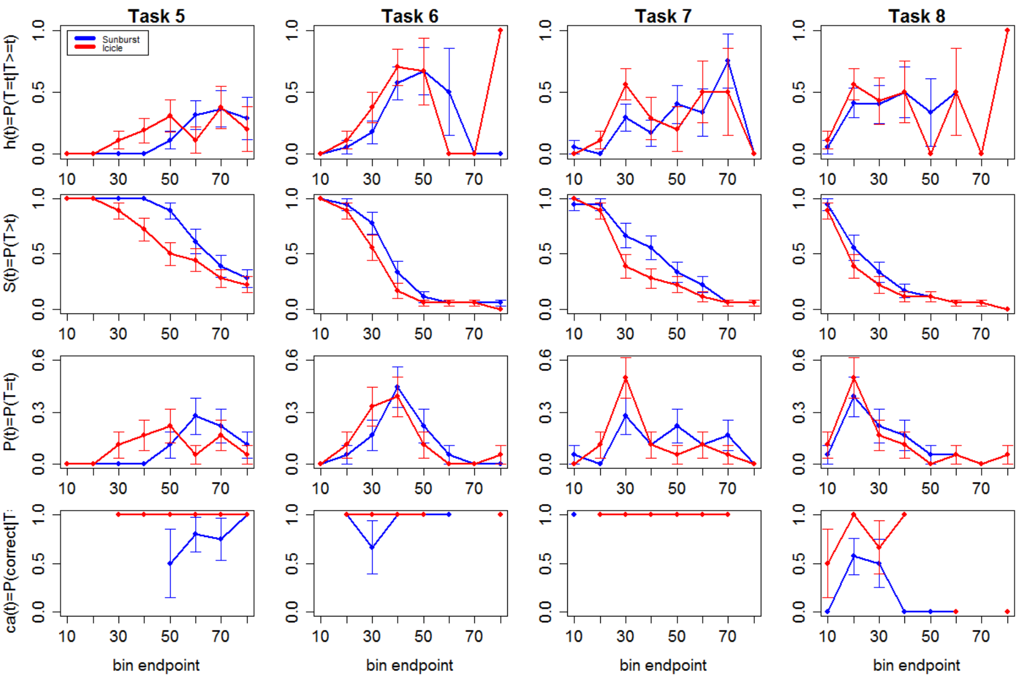

Figure 6. Sample-based descriptive statistics for Tasks 5 to 8 in the control study setting. Hazard estimates h(t), survival estimates S(t), (sub)probability mass estimates P(t), and conditional accuracy estimates ca(t) for the both layouts are presented in rows 1, 2, 3, and 4 respectively

4.3 Discussion

Earlier, we hypothesized that the participants will show better performance with the new Icicle layout for effectiveness and efficiency metrics. Our results support these hypotheses.

To summarize, we first conducted our study under the controlled study setting followed by the online study setting. In both study settings, we compared two groups, with either the Sunburst layout or the Icicle layout. The data for the controlled setting were obtained through a lab experiment, while those for the online setting were obtained through online forms. We used two different research methodologies to analyze the data: the comparison of mean RT and mean accuracy using standard t-test, and the event history analysis. The mean accuracy and the response time measures showed a trend that favored the Icicle layout. The subjective questionnaire responses showed a strong preference for the Sunburst layout in both settings, the controlled and the online. The event history analysis showed that participants were not only quick to respond in most of the tasks with the Icicle layout but also that they were accurate. The analysis also showed that for the two tasks, when participants responded quicker with the Sunburst layout, there was a low probability for the answers to be accurate. This finding suggests that not only were the participants slower in most of the tasks with the Sunburst layout, but they were also inaccurate when responding quickly. Furthermore, through the hazard plots we were able to show that participants were slower with those tasks that involved file comparisons (for example Tasks 2 and 4). Similarly, we also showed that participants were quicker with the directory-based comparison tasks (for example Tasks 1, 3, and 8).

From the subjective responses for the layout preference, we observed that the participants preferred the Sunburst layout over the Icicle layout; although participants were faster and more accurate with the Icicle layout. While we are unsure of the reason, we can say that the participants' preference for a particular layout is independent from their performance while they were working on it. As predicted in past research studies, we hypothesized that user acceptance for the Sunburst layout will be greater than the Icicle layout. In our study, we replicated this finding in both controlled and online study settings (refer to Table 3 and 4). However, the difference was clearer in the case of the online study setting. One possible reason for such a clear difference in the online study setting could be because of the user friendly environment, as it is more convenient for the participants to interact and provide ratings in their own preferred environment. In the controlled setting, this is not the case, as participants were more likely to give ratings under the controlled supervision of one or more moderators in the laboratory setting.

Overall, we successfully showed how layout impacts user performance. Using the different analysis approaches, we showed different perspectives on the same data. By using the event history analysis approach, we were able to conclude that measuring only means and comparing them between groups may not provide the detailed information we require for evaluation studies.

5. CONCLUSION

In our studies, we evaluated two display space-filling techniques, i.e. the Sunburst (radial) layout and the Icicle (rectangular) layout. By conducting a controlled study and an online study, we showed the usability of the two visualization tools in terms of accuracy, efficiency and user acceptance.

From our results, we conclude that the accuracy and efficiency results favored the Icicle layout while the user acceptance results favored the Sunburst layout. Through EHA methodology, we were able to better characterize the temporal dynamics of user performance in our tasks compared to the mean response times (RT) analysis. Although, the mean RT analysis was able to demonstrate trends in favor of the Icicle layout, there was more concrete evidence for the latency-specific differences between the layouts from the EHA results. Through this, we successfully demonstrated how time-to-event data can be useful to reveal insights that are concealed by the mean.

Future scope for this research extends to screen recording analysis. Using screen recording software, we can record participants' actions while performing the tasks. This way we can closely watch their actions and clicks and track the path taken to reach their final conclusions during solving of the task

REFERENCES

1. Abdullah, N., Adnan, W. A. W., and Noor, N. L. M. 2011. Cognitive Design of User Interface: Incorporating Cognitive Styles into Format, Structure and Representation Dimensions. *In Digital Information and Communication Technology and Its Applications*, pp. 743-758, Springer Berlin Heidelberg.
2. Allison, P. D. 1982. Discrete-Time Methods for the Analysis of Event Histories. *Sociological Methodology*, vol. 13(1), pp. 61-98.
3. Allison, P. D. 2010. *Survival Analysis Using SAS: a practical guide.* Sas Institute.
4. AlTarawneh, R., Humayoun, S. R., and Al-Jaafreh, A. 2015. Towards Optimizing the Sunburst Visualization for Smart Mobile Devices. *In NTERACT 2015 Adjunct Proceedings: 15th IFIP TC. 13 International Conference on Human-Computer Interaction*, 14-18 September 2015, Bamberg, Germany, pp. 323-333, University of Bamberg Press.
5. AlTarawneh, R. and Humayoun, S. R. 2016. Visualizing Software Structures through Enhanced Interactive Sunburst Layout. *In Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI '16)*, Paolo Buono, Rosa Lanzilotti, and Maristella Matera (Eds.). ACM, New York, NY, USA, pp. 288-289.
6. Callaway, C., Kuflik, T., Not, E., Novello, A., Stock, O., and Zancanaro, M. 2005. Personal Reporting of a Museum Visit as an Entrypoint to Future Cultural Experience. *In Proceedings of the 10th international conference on Intelligent user interfaces (IUI '05)*, ACM, New York, NY, USA, pp. 275-277.
7. Chechile, R. A. 2003. Mathematical Tools for Hazard Function Analysis. *Journal of Mathematical Psychology*, vol. 47(5), pp. 478-494.
8. González-Torres, A., García-Peñalvo, F. J., and Therón, R. 2013. Human–Computer Interaction in Evolutionary Visual Software Analytics. *Computers in Human Behavior*, vol. 29(2), pp. 486-495.
9. González Torres, A., García Peñalvo, F. J., and Therón Sánchez, R. 2016. How Evolutionary Visual Software Analytics Supports Knowledge Discovery. *Science of Computer Programming, Special Issue on Knowledge-based Software Engineering*, vol. 121, 1 June 2016, pp. 55–74.
10. Greenhouse, J. B., Stangl, D., & Bromberg, J. 1989. An Introduction to Survival Analysis: Statistical Methods for Analysis of Clinical Trial Data. *Journal of Consulting and Clinical Psychology*, vol. 57(4), pp. 536-544.
11. Herman, I., Melanon, G., and Marshall, M.S. 2000. Graph Visualization and Navigation in Information Visualization: A Survey. *IEEE Transaction on Visualization and Computer Graphics*, vol. 6(1), pp. 24–43.
12. Johnson, B. and Shneiderman, B. 1991. Tree-Maps: A Space-Filling approach to the Visualization of Hierarchical Information Structures. *In Proceedings of the 2nd Conference on Visualization '91*, pp. 284–291, Los Alamitos, CA, USA, IEEE Computer Society Press.
13. Katrina L. P. 2012. Information Literacy on the Web: How Visual and Textual Cues Contribute to Website Credibility Assessments. *Communication in Information Literacy*, vol. 6(1), pp. 34-48.
14. Keiley, M. K., and Martin, N. C. 2005. Survival Analysis in Family Research. *Journal of Family Psychology*, vol. 19(1), pp. 142-156.

15. Luce, R. D. 1986. Response Times: *Their Role in Inferring Elementary Mental Organization*. Oxford University Press on Demand.
16. Morita, J. G., Lee, T. W., & Mowday, R. T. (1989). Introducing survival analysis to organizational researchers: A selected application to turnover research. Journal of Applied Psychology, 74(2), 280.
17. Núñez-Antón, V., & Orbe, J. (2005). Statistical Time to Event Analysis in the Social Sciences: Modeling Hazard Rate and Duration in Finance. Methodology, 1(3), 104-118.
18. Oviatt, S. 2006. Human-Centered Design Meets Cognitive Load Theory: Designing Interfaces that Help People Think. *In Proceedings of the 14th ACM international conference on Multimedia (MM '06)*, pp. 871-880, ACM.
19. Panis, S., and Hermens, F. 2014. Time Course of Spatial Contextual Interference: Event History Analyses of Simultaneous Masking by Nonoverlapping Patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 40(1), pp. 129-144.
20. Panis, S., and Schmidt, T. 2016. What is Shaping RT and Accuracy Distributions? Active and Selective Response Inhibition Causes the Negative Compatibility Effect. *Journal of Cognitive Neuroscience*, vol. 28(11), pp. 1651-1671.
21. Proetzsch, M., Luksch, T., and Berns, K. 2010. Development of Complex Robotic Systems Using the Behavior-based Control Architecture iB2C. *Robotics and Autonomous Systems*, vol. 58(1), pp. 46–67.
22. Rice, W. R. 1989. Analyzing Tables of Statistical Tests. *Evolution*, vol. 43(1), pp. 223-225.
23. Singer, J. D., and Willett, J. B. 2003. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence.* Oxford University Press.
24. Stasko, J., Catrambone, R., Guzdial, M., and McDonald, K. 2000. An Evaluation of Space-Filling Information Visualizations for Depicting Hierarchical structures. *International Journal of Human-Computer Studies*, vol. 53(5), pp. 663–694.
25. Stasko, J. and Zhang, E. 2000. Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations. *In Proceedings of the IEEE Symposium on Information Visualization (INFOVIS '00)*, pp. 57–, Washington, DC, USA, IEEE Computer Society.
26. Turo, D., and Johnson, B. 1992. Improving the Visualization of Hierarchies with Treemaps: Design Issues and Experimentation. *In Proceedings of the 3rd conference on Visualization '92*, pp. 124-131, IEEE Computer Society Press.
27. Willett, J. B., and Singer, J. D. 1995. It's Déjà Vu All Over Again: Using Multiple-Spell Discrete-Time Survival Analysis. *Journal of Educational and Behavioral Statistics*, vol. 20(1), pp. 41-67.