Survey on User Studies on the Effectiveness of Treemaps

Carolin Fiedler
Hasso Plattner Institute,
Faculty of Digital Engineering,
University of Potsdam,
Germany

Willy Scheibel
Hasso Plattner Institute,
Faculty of Digital Engineering,
University of Potsdam,
Germany

Daniel Limberger
Hasso Plattner Institute,
Faculty of Digital Engineering,
University of Potsdam,
Germany

Matthias Trapp
Hasso Plattner Institute,
Faculty of Digital Engineering,
University of Potsdam,
Germany

Jürgen Döllner
Hasso Plattner Institute,
Faculty of Digital Engineering,
University of Potsdam,
Germany

ABSTRACT

Treemaps are a commonly used tool for the visual display and communication of tree-structured, multi-variate data. In order to confidently know when and how treemaps can best be applied, the research community uses usability studies and controlled experiments to "understand the potential and limitations of our tools" (Plaisant, 2004). To support the communities' understanding and usage of treemaps, this survey provides a comprehensive review and detailed overview of 69 user studies related to treemaps. However, due to pitfalls and shortcomings in design, conduct, and reporting of the user studies, there is little that can be reliably derived or accepted as a generalized statement. Fundamental open questions include configuration, compatible tasks, use cases, and perceptual characteristics of treemaps. The reliability of findings and statements is discussed and common pitfalls of treemap user studies are identified.

CCS CONCEPTS

• **Human-centered computing** → **Treemaps**; **Visualization design and evaluation methods**.

KEYWORDS

Treemap, Evaluation, Literature Review

1 INTRODUCTION

For much of today's tree-structured data, treemaps are a versatile and established tool for visualization [37]. Since the introduction of treemaps [25], more than 400 publications have been published on this topic. As for any visualization, it is essential to "show that [it] is useful in a practical sense" [28] and "to understand the potential and limitations" [42] of it. This work examines 69 related user studies from 1992 to 2018 and reviews to what extent these "provide convincing evidence of utility" [42]. In many of these studies, mistakes in preparation or implementation have been made that led to exaggerated statements. As a consequence, "it is common for claims to be made that go far beyond the results" and "many of the studies we read are biased" [55]. Using the available data, it is difficult to decide which visualization is best suited for a particular task [29].

With this paper, we contribute a review of these user studies by means of overview tables. Using the detailed data on the user studies, we conduct an analysis of common pitfalls and discuss possible mitigation. The full data set is available by means of an interactive community website and public GitHub repository[1]. The remainder of this paper is structured as follows. Section 2 recapitulates the basics of user studies in the field of information visualization. Section 3 reviews available user studies on treemaps. Section 4 provides a discussion on the findings and quality of the user studies. In Section 5, we summarize user study guidelines and publication best-practices. Section 6 concludes this paper.

2 RELATED WORK AND METHODS

Evaluations can examine various aspects of the transformation from data to visual artifacts and the perceptual characteristics of those. They provide the opportunity to detect shortcomings and pitfalls but also reveal suitable applications and the potential of particular parameterization. Common evaluation methods are metric-based evaluations, heuristics, or case studies. User studies are a popular method as they "offer a scientifically sound method to measure a visualization's performance" and "can be used to evaluate the strengths and weaknesses of different visualization techniques" [28]. Further, user studies are able to provide data on which approach suits a certain objective best and which properties are more likely to achieve desired results. It investigates "why a particular technique is effective" and whether it is "useful in a practical sense" [28].

Similar to this paper, other publications reviewed a body of work to assess user studies and other means of evaluation for visualization techniques. As main distinction, these reviews do not focus

[1]github.com/varg-dev/treemap-studies

on treemap visualization techniques but filter by publication venue or problem domain. As such, Lam et al. tagged 361 papers on user studies in information visualization and derived seven scenario categories [29]. These categories were extended by reviewing 581 user studies [23]. Resulting from this, the authors concluded that the research community focuses on "evaluations of algorithmic performance […] and qualitative result inspections […] through images" [23] with a lack on *communication through visualization* and *collaborative data analysis* for their initial set of publications. With a focus on the software visualization domain, Merino et al. provided a review on software evaluation and visualization techniques covering 181 evaluation papers, whereof 81 papers documented evaluation with user feedback [38].

2.1 Methods for User Evaluation

Users can participate in various ways in an evaluation process and, depending on the type of participation, provide valuable insights into various aspects of a visualization. Researchers have to choose suitable evaluation techniques depending on their intentions and available means. The different aspects of evaluation result in multiple independent categorizations of user studies. Consequently, it is advisable to consciously determine methodology and content for carrying out an evaluation, as proposed by Leroy et al. [30] and Plaisant [42], respectively. Alternative categorizations for information visualization evaluation were presented by Lam et al. [29] and its revision by Isenberg et al. [23], which offer a more holistic approach. As an appropriate analysis would go beyond the scope of this report, this review concentrates on the former two, more focused categorizations.

Types of User Studies for Software Evaluation. Gondy Leroy suggested a summary of most common evaluation types for software evaluation [30]. Although researchers may use *case studies*, *field studies*, or *descriptive studies* to explain and answer difficult questions, they are not necessarily representative [42]. To obtain reproducible results, controlled experiments or their moderated variants are needed:

Correlation studies describe changes in certain variables. The intention is to identify where a change in one variable coincides with change in other. They cannot detect the direction of causation though. *Controlled experiments* evaluate the impact, benefit or drawbacks of information systems. They are often referred to as demonstration studies, randomized controlled trials, or comparative studies. Researchers control one or more variables to measure values of one or multiple outcomes. This allows to validate causation. If a random assignment of participants to experiment conditions is not possible—for example due to time, geographical, or social constraints—a *quasi experiment* is conducted instead. The chosen type defines the transferability of the user study results to a more general level and whether causation could be shown.

Evaluation Practices. Catherine Plaisant [42] presented a categorization for the user evaluation of information visualization. The four categories provide insights on what is compared in which environment (comparison method). First, *controlled experiments comparing different techniques* are user studies that compare two or more different techniques. These are referred to as *visualization*

techniques comparisons (VTC) and are typically used to contrast new techniques with existing ones. The second category is *controlled experiment comparing design elements* of an individual visualization technique. Such studies compare a parameterization with another configuration, e.g., mappings of information to visual display. For treemaps, we consider three sub-categories: *layout*, *visual variables*, and *configuration* of treemap. The third category is *usability evaluation*. It can provide feedback on a certain visualization tool and prioritize its subsequent improvement and refinement. Finally and least common, *case studies* enable evaluation in realistic settings, and, similar to Leroy's categorization, facilitate to demonstrate in-context usefulness. The results of a case study, however, may not be replicable and generalizable [42]. Such studies are therefore not included. During our review we found a number of user studies that do not fit into the categories. They all deal with treemaps on a meta level in relation to communication and aesthetic perception, or more precisely visual quality, and are assigned to a *meta* category.

2.2 Design and Conduct of User Studies

If researchers want to prove a causation, they have to conduct a controlled experiment. Especially for such experiments, possible results can become unusable because of smallest errors. Mistakes often occur during the preparation or execution of a user study that result in incorrect statements. Published results of erroneous user studies hardly represent added value for the research community. Regarding this, Colin Ware states that "it is common for claims to be made that go far beyond the results" [55], e.g., generalizing satisfaction statements from single participants [23]. Researchers are often unaware of the specifics of those errors and mistakes in the preparation and conduct of user studies. Consequently, "many of the studies we read are biased" [55]. The main objective should therefore be to conduct informative, value-adding and qualitatively meaningful user studies that meet critical validation priorities.

Design problems often emerge from a discrepancy between the research goal and the experiment actually conducted. Typical, critical design decisions are determining the independent variables and tasks participants have to perform. Further difficulties are the formulation of hypotheses that fit the research goal, the establishment of appropriate control conditions, and the occurrence of interfering variables [55]. In the broad field of statistical analysis, researchers might lose track or are not aware of implications and exact meaning of results. Despite methodological sources of error, some user studies lack appropriate and fair control conditions—e.g., another visualization that does not implement a feature in a comparable way—and cause an a-priori biased result. On top of that, the expectation of the research community has an influence on the way the results of an experiment are published: it "is easier to publish results that confirm a hypothesis than results showing no effects" [55]. However, this promotes claims "that go far beyond the results" and seems to be a common, though certainly unintended, practice.

2.3 What We Should Expect from User Studies

The evidence of utility and a deep knowledge of tasks or use cases in relation to different visualization techniques are of major importance. Well thought-out concepts, careful conduct, and fair control

conditions are required. Researchers should (1) guarantee comprehensibility of their studies by describing their thoughts and procedure precisely, (2) publish their data to support validation and verification of their results, as well as (3) formulate results honestly and distinguish clearly from claims and assumptions.

Study replications are essential to confirm results in a validation process but those are rare [23, 55]. In addition, researchers should evaluate not only their tool, but also independently evaluate the underlying visualization technique and used technology. For this, however, it is essential to understand any implementation of a visualization as a specialized instance of that visualization (not the visualization itself). The utility of a specific treemap implementation, for example, might be strongly effected by its aesthetics, responsiveness, colors and shading, size, use of additional graphical effects, embedding in a GUI and the GUI itself, peripherals, and many other aspects. These, however, are aspects that are not necessarily an inherent part of a treemap visualization technique. If a sound distinction or exclusion of such aspects is not feasible, the goals of respective studies should be reviewed exhaustively and awareness of respective factors should be increased.

In order to assess the power of treemaps, derive qualified statements from previous results, and judge the research area—regarding topic coverage, clusters, and outliers—a model assessing visualization techniques from multiple usage perspectives is helpful. Since there are no widely used models on evaluation topics in information visualization yet [65], we propose the following set:

Use Cases refer to specific tasks, data domains, data structures, and research questions subject to the visualization technique.
Perception refers to human perception and dependencies within visual encodings.
**Interaction** refers to the interaction metaphors applied and their technical implementations and constraints.
Communication refers to the language and metaphors used to introduce and talk about visualization techniques and tasks.
Configuration refers to specific choice of values for a parameterized visualization technique.

These concepts relates as follows. To solve a *communicated* task within an *use case* scenario, a user *perceives* the visual output of a properly *configured* information visualization and *interacts* with it.

3 USER STUDIES ON TREEMAPS

The evaluation of treemaps started almost simultaneously with their invention in 1991 [24, 49]. In contrast to the initial definition of treemaps as a recursive subdivision of a 2D surface [25], hereinafter treemaps denote visualizations of data based on parent-child-relations where child nodes are located within their parent surfaces but are not required to utilize all of the available parent node's space [5]. Thereby, treemaps include not only subdivisions of 2D rectangles but the extension to *n*-dimensional spaces [24] and other shapes [2, 3, 56]. Since the initial publication in 1991, over 400 publications regarding treemaps were published in conference proceedings and journal articles of respective publishers, i.e., IEEE (TVCG, iV, VISSOFT, VIS, INFOVIS, and VAST), ACM (ToG, VINCI, OzCHI, ICSE, and AVI), Springer (CompStat, ISWC, ITS, and APCHI), EG (CGF, EuroVis, and EuroRVVV), SciTePress
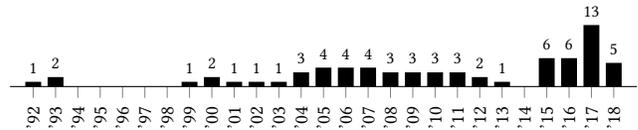
Figure 1: Number of user studies considered per year. The publication year 2018 is not completely considered.

(VISIGRAPP / IVAPP), and SagePub (IVI). For the list of papers considered in this review, we extended this corpus by (1) selected doctoral theses and technical reports as some user studies were not published with peer review, (2) a short list of individual papers by lesser frequented publishers, their journals, and conferences. We selected papers that include evaluations and studies with participants, whereby the study is classifiable as *Summative User Study (Evaluation)* [41]. If multiple user studies are documented within one publication, they are treated as different studies within the review and analysis. We identified 69 user studies in according publications on treemaps until 2018 and considered all of them for structural analysis and 63 of them for the evaluation of the results. Thereby, we excluded case studies as they are hard to reproduce and the statement generalization is limited [42]. More details on the excluded studies is provided in Section 3.2. The per-year distribution is shown in Figure 1.

3.1 Overview of User Studies

The results of the survey are provided as a comprehensive data set of which an excerpt is visualized in Table 1 and Table 2. Table 3 contains information on compared visualization techniques. The complete data set, its schema, and a description for every data category is interactively available on the community website[1] on which researchers are invited to collaborate as well as refine and extend the data set.

In the following, the user studies are described statistically. Unless otherwise stated, all figures refer to all 69 user studies analyzed: of the user studies analyzed, almost half contain comparisons with other visualization techniques (*VTC*) and more than one quarter *configuration* aspects. There are only three studies targeting *layout*, ten studies evaluating *visual variables*, and four studies in the *meta* category. In *VTC* studies, typically multiple visualizations are compared. Comparisons with node-link diagrams were the most frequent, followed by sunburst views and lists. Of the six layout algorithms, only the *Strip* and *Squarified* algorithms appeared in all layout user studies. The evaluation data type was purely *qualitative* in four studies and purely *quantitative* in more than three quarters of the studies. The latter use more than 80 % of *effectiveness* and *efficiency* measures [30] for their dependent variables. Less than half used *satisfaction* measures. The size of used data sets ranged broadly from five to 200 000 nodes with *VTC* studies using the most extensive data sets in average and *meta* studies the smallest. For studies targeting *visual variables*, most did not include an actual real-world data set. Almost three quarters of the studies were conducted in a *laboratory* setting, while *online* and *Mechanical Turk* studies accounted for only one eighth each. The number of participants ranged from five to 285 (Figure 2).

Table 1: Excerpt of treemap survey, part one—including an assessment on study categorization, specific evaluation details, characteristics of the used data set, and chosen treemap configuration (●).
Notation: evaluation with (★) / without (◇) statistically significant results reg. treemaps │ ▪ cell not applicable │ – unknown
Abbreviations: Dimension (*Dim.*) │ Implementation (*Impl.*) │ Topology Depiction (*Topology*) │
Visualization Technique Comparison (*VTC*) │ Visual (*Vis.*)

| | Category | | | | | | Evaluation | | | | | | | | | | | | | Dataset(s) | | | | | | | | | | | Treemap Configuration | | | | | | | | | |
| | Evaluation Practice | | | | | | Data | | Impl. | | | #Participants | | | | | Measure | | | Topology | | | | Origin | | | Size (#n) | | | | Shape | | | Dim. | | | Topology | | | |
	VTC	Design: Layout	Design: Vis. Variable	Design: Configuration	Usability	(Meta)	quantitative	qualitative	laboratory	online	MTurk	<10	10–29	30–49	50–100	>100	effectiveness	efficiency	satisfaction	list	hierarchy	tree	graph	artifical	adapted	real-world	<10	10–100	100–1000	>1000	rectangular	circular	map-like	2D	2.5D	3D	non-nested	non-nested: cushion	nested	nested: titlebar
Directory Browsing [24, 49]	★						●		●				●				★	★	★			●				●		●			●			●					●	
GEnie Experiment [49]	★		◇				●		●				●					★				●			●	●		●			●			●						●
US Budget Experiment [24]	★						●		●					●			★	★	◇			●			●	●		●			●			●					●	
3D Information Visualizations [58]	★				★		●		●				●				★	★	◇			●		●		●		●			●					●			●	
Sunburst vs. Treemap (small) [46]	★						●	●	●					●			★	◇	◇			●			●			●			●			●			●			
Sunburst vs. Treemap (large) [46]	★						●	●	●					●			◇	◇	★			●			●				●		●			●			●			
Topology & Node Size [4]	★						●		●				●				★	★	★			●		●		●		●	●		●			●					●	
Compare to Beamtrees [50]	★						●		●				●				★	★				●		●			●	●			●			●				●	●	
Compare 5 Tree Visualizations [26]	★			★	★		●	●	●					●			★	★	★			●				●		●			●			●				●		●
Treemaps & Tables [16]	★	★					●		●				●				★	★				●			●	●		●			●			●				●	●	
Crop Circles [53]	★			★	★		●		●				●				★	★	◇	●					●			●			●	●		●						●
Knowledge Discovery [54]	★						●		●				●				★	★	★			●			●					●	●			●						●
Hierarchy Browser [1]	★		◇				●		●					●			★	★				●			●				●		●			●						●
Search Engine Results [12]	◇						●	●		●		●					◇	◇				●			●		●			●				●						●
Multiple Views [47]	★			★				●	●				●									●			●		●			●			●					●	
Search Result Interface 1 [13]	★						●		●				●				◇	◇	★	●					●		●			●			●					●	
Search Result Interface 2 [13]	★						●		●					●			◇	◇	★	●					●		●	●		●			●					●	
Data Density [27]	★						●			●			●				★	★				●		●			●	●		●			●				●		
Node Link Treemaps [35]	◇						●		●				●				◇	◇				●			●			●			●				2.5D:●				●	
Software Systems as Cities [57]	●						●		●					●			★	★	◇			●			●			●			●				2.5D:●				●	
Eye-tracking for EER-Tutor [36]	★						●	●	●				●				★	★		–	–	–	–		●	●		●			●			●					●	
Rapid Overview [39]	★						●	●	●				●				★	◇				●			●			●	●		●			●					●	
Software Landscape [15]	★						●	●	●				●				◇	★	◇			●			●			●	●		●			●					●	
Treemap vs. Euler [43]	◇						●		●				●				◇	◇			●	●			●			●			●			●					●	
Treemap vs. Euler 2 [44]	◇						●		●				●				◇	◇			●	●			●			●			●			●					●	
Photo: Layout [18]	★						●				●					●	★	★	★	●					●			●			●			●				●		
Compare 3 Tree Visualizations [40]	★			★			●	●	●							●	★	◇				●		●			●	●			●			●						●
Wrapped Bars: Comparison [60]	★						●			●			●				★	◇		●					●			●	●		●			●				●		
Wrapped Bars: Ranking [60]	★						●			●				●			★	★		●					●			●	●		●			●				●		
Wrapped Bars: Distribution [60]	★						●			●			●				★	★		●					●			●	●		●			●				●		
Pie Charts: Proportion [21]	★						●		●				●				★	★		●					●			●			●			●				●		
Pie Charts: Comparison [21]	★						●		●				●				★	★		●					●			●			●			●				●		
Pie Charts: Labeling [21]	★						●		●				●				★	★		●					●			●			●			●				●		
Pie Charts & Stacked Bars [21]	★						●		●				●				★	◇	◇	●					●			●			●			●				●		
Compare Layout Algorithms [5]		★					●		●				●				★	★	★	●					●			●			●			●				●		
Relative Direction Change [19]		★					●		●				●					★			●				●		●	●		●			●					●	
Layout in Cartography [61]		★					●			●					●		★	★			●				●		●	●		●			●					●	
Luminance (Pilot) [27]			★				●			●				●			★				●		●			●			●			●			●				●		
Aspect Ratio [27]			★				●			●			●				★			▪	▪	▪	▪								●			●			▪	▪	▪	▪
Rectangular Orientation [31]			★				●		●				●				★	◇		▪	▪	▪	▪								●			●			▪	▪	▪	▪
Sketchiness: Perceived Height [32]			★				●			●					●		★														●				2.5D:●					
Sketchiness: Pre-Attentiveness [32]			★				●		●				●				★				●		●			●	●			●				2.5D:●			●		
Sketchiness: Order [32]			★				●		●					●			★			▪	▪	▪	▪								●				2.5D:●		▪	▪	▪	▪
Color Scheme: Model [59]			★				●		●				●				★			●					●			●			●			●				●		
Color Scheme: Aesthetics [59]			★				●		●					●			★			●					●		●	●		●			●				●		
Color Scheme: Size [59]			★				●		●					●			★	★		●					●		●	●		●			●				●		
Color Scheme: Combination [59]			★				●		●					●			★			●					●		●	●		●			●				●		
2D and 2.5D [8]				★			●		●				●				★	★			●			–	–	–				●			●		●	●			●	
Animation in 2.5D [7]				◇			●	●	●				●				◇	◇			●			–	–	–				●			●			●			●	
Content Browsing 1 [45]				★			●		●				●					★			●				●			●			●				2.5D:●				●	
Content Browsing 2 [45]				★			●		●				●				◇	★			●				●			●			●				2.5D:●				●	
Cushion Treemaps [22]				★			●		●				●				★	★	◇		●				●			●			●			●				●	●	
Cushion Treemaps (Structure) [22]				★			●		●				●				◇	★			●				●			●			●			●				●	●	
Image Browser [17]				★			●		●				●					★	◇		●			–	–	–				●			●					●	
Balloon Focus [48]				★			●		●				●				★	★	★		●				●			●			●			●						●
Photo: Count [18]				★			●				●					●	★	★	★	●					●			●			●			●				●		
Photo: Flickr [18]				★			●			●		●						★	★	●					●			●	●		●			●				●		
Treemaps vs. Map-like [6]				★			●		●				●				★	★			●				●	●			●		●		●	●			▪	▪	▪	▪
Aggregation: Visual Search [33]				★			●			●			●				★	◇	★		●				●				●		●				2.5D:●			●		
Aggregation: Readability [33]				◇			●			●			●				◇				●				●				●	●	●				2.5D:●			●		
Height References [34]				★			●			●			●				★	◇			●				●			●			●								●	
Flexible Hierarchy Interface [11]					★			●	●			●					▪	▪	▪		●				●			●	●		●			●						●
Virtual Reality [14]					★		●	●	●				●				▪	▪	▪			●				●		●	●		●					●				●
Code Maps 1 [20]					★			●	●			●										●				●			●	●	●			●	2.5D:●			●		●
Code Maps 2 [20]					★			●	●			●										●				●				●			●	2.5D:●			●		●
Aesthetic [10]						★	●			●						●	◇	★	★			●				●		●	●		●			●	2.5D:●			●		●
Visual Metaphors [63]						★	●		●					●			★	★				●		●				●		●			●						●	●
Individual Visual Metaphors [64]						★	●				●			●		●	★	★	★			●				●			●		●			●						●
LoC & Visual Metaphors [62]						★	●				●					●	★	★	★			●				●		●	●		●			●						●

Table 2: Excerpt of treemap survey, part two—including an assessment on details on chosen treemap configuration, interaction, user interface, and study tasks (•).

Notation: ■ table cell not applicable | – unknown |

Abbreviations: Verbal Metaphor (*Mtph.*) | Dimension (*Dim.*) | size (*S*) | height (*H*) | color (*C*) | structure (*R*) | (sub)tree-attribute (*T*) | count (*U*) | name/id (*N*) | locate (*L*) | other attribute/visual variable (✦).

Study	Offset: none	contour	thin	broad	Label: none	categories	leaves	full	Nav: none	node zooming	restricted	free	UI: none	search	filtering	attribute mapping	mapping parameters	adjustable offset	get attribute values	identify value	access information	sort	compare	identify distribution	distinguish	associate	correlate	create information	other	containment	levels	1D	2D	3D
Directory Browsing [24, 49]			•					•	•								•	•	•	NSU✦	LCR	S			S					•		•	•	
GEnie Experiment [49]	•								•	•									•	–		–	–	–	–	–	–	–	–		–	–	–	–
US Budget Experiment [24]				•				•		•							•	•	•	NSU	LSR	S	S		S							•	•	•
3D Information Visualizations [58]			•					•			•	•								U	L	R								•		•		
Sunburst vs. Treemap (small) [46]		•					•		•										•		LSC	SR			SR	✦				•		•	•	•
Sunburst vs. Treemap (large) [46]		•					•		•										•		LSC	SR			SR	✦				•		•	•	•
Topology & Node Size [4]				•				•	•				•							T	LR	S			SR						•	•		
Compare to Beamtrees [50]	•	•		•				•		•			–	–	–	–	–	–	–	NT✦	LR	S			SR		✦			•		•	•	•
Compare 5 Tree Visualizations [26]	•		•					•		•					•	•	•	•	•	NUT	LCR✦	T			SR✦					•		•	•	•
Treemaps & Tables [16]			•					•	•	•							•		•	U✦	LS	R			S✦	✦	✦			•		•	•	
Crop Circles [53]		•		•				•	•	•			•		•				•		LR	R			R			✦				•	•	
Knowledge Discovery [54]		•						•		•					•	•	•	•	•	N						✦								
Hierarchy Browser [1]		•						•	•				•							NU	LR	R			R					•		•		
Search Engine Results [12]				•				•	•								•	•		–	–	–	–	–	–	–	–	–	–					
Multiple Views [47]				•				•	•				•															✦				•	•	
Search Result Interface 1 [13]			•				•		•				•							–	–	–	–	–	–	–	–	–	–			–	–	–
Search Result Interface 2 [13]			•				•		•				•							–	–	–	–	–	–	–	–	–	–			–	–	–
Data Density [27]		•						•	•				•									S	S									•		
Node Link Treemaps [35]	•		•									•							•	NUT	LR✦	T			SR							•	•	
Software Systems as Cities [57]			•	•							•				•	•		•	•		C✦	✦			HR✦		✦	✦				•	•	•
Eye-tracking for EER-Tutor [36]		•						•	•				•							–	–	–	–	–	–	–	–	–	–		–	–	–	–
Rapid Overview [39]		•						•	•				•							U	R	S						✦				•	•	
Software Landscape [15]				•				•	•				•							N									✦					
Treemap vs. Euler [43]		•					•		•				•							–	–	–	–	–	–	–	–	–	–			•		
Treemap vs. Euler 2 [44]		•					•		•				•							–	–	–	–	–	–	–	–	–	–			•		
Photo: Layout [18]	•			•					•				•								L				S							•		
Compare 3 Tree Visualizations [40]		•						•	•				•							U	R	S						✦			•	•		
Wrapped Bars: Comparison [60]	•			•				•	•				•									S	S									•		
Wrapped Bars: Ranking [60]	•			•				•	•				•									S										•		
Wrapped Bars: Distribution [60]	•			•				•	•				•											S								•		
Pie Charts: Proportion [21]	•						•		•				•							S												•		
Pie Charts: Comparison [21]	•						•		•				•										S									•		
Pie Charts: Labeling [21]	•						•		•				•										S									•		
Pie Charts & Stacked Bars [21]	•						•		•				•										S									•		
Compare Layout Algorithms [5]	•						•		•				•								L											•		
Relative Direction Change [19]			•				•		•				•								L											•		
Layout in Cartography [61]			•				•	•					•							N	LR	S		SC			✦	✦		•	•	•		
Luminance (Pilot) [27]	•			•				•	•				•									S	S									•		
Aspect Ratio [27]																						S	S											
Rectangular Orientation [31]																						S	S											
Sketchiness: Perceived Height [32]									•																H							•		
Sketchiness: Pre-Attentiveness [32]			•					•	•					•												✦						•		
Sketchiness: Order [32]																						✦										•		
Color Scheme: Model [59]																						S										•		
Color Scheme: Aesthetics [59]	•							•					•																✦		✦			
Color Scheme: Size [59]	•								•				•									S										•		
Color Scheme: Combination [59]	•							•	•				•									S							✦		✦	•		
2D and 2.5D [8]		•					•				•					•			•	N	LCR	ST		C	SR					•		•	•	
Animation in 2.5D [7]		•					•			•	•					•	•				LCR			C	S			✦		•		•	•	
Content Browsing 1 [45]		•					•		•				•								L											–	–	
Content Browsing 2 [45]		•					•		•				•								L											–	–	
Cushion Treemaps [22]		•	•				•			•			•							NU	LCR			C	SR					•		•	–	
Cushion Treemaps (Structure) [22]	•	•					•		•				•															✦			✦	•	–	
Image Browser [17]			•				•				•	•									L											•		
Balloon Focus [48]			•				•		•								•		•	U	C✦				S			✦			✦	•		
Photo: Count [18]	•						•		•				•								L										✦			
Photo: Flickr [18]	•						•	•	•				•															✦			✦	–	–	–
Treemaps vs. Map-like [6]			•					•	•				•																	•		•		
Aggregation: Visual Search [33]		•					•		•																S✦	✦				•		•	•	
Aggregation: Readability [33]		•					•		•				•							✦						✦				•		•	•	
Height References [34]			•					•	•							•									H	✦						•		
Flexible Hierarchy Interface [11]			•					•	•						•	•	•	•	•	N✦	✦		S		S	✦						•	•	
Virtual Reality [14]			•					•	•		•	•								N	LR				S		✦	✦		•		•	•	
Code Maps 1 [20]	•						•		•				•													✦		✦			✦			
Code Maps 2 [20]	•											•	•		•						L							✦			✦			
Aesthetic [10]	•	•			–	–	–	–	•				•							–	–	–	–	–	–	–	–	–	–			–	–	–
Visual Metaphors [63]		•						•	•	•			•							NU	LCR				R					•	•	•	•	•
Individual Visual Metaphors [64]		•						•	•				•								LR	RT									•	•	•		
LoC & Visual Metaphors [62]				•				•	•	•			•								LS✦					✦		✦		•	•	•	•	•

The most common configuration of treemaps were *2D rectangular* treemaps. More than 80 % of the studies included rectangular treemaps, while *circular* and *map-like* shapes were uncommon, with one and three studies, respectively. There were only one study covering *3D* treemaps and 12 studies containing *2.5D* treemaps. More than half of the studies included *non-nested* treemaps [24] (including *cushion* treemaps [51]) and almost 40 % of the studies *nested* treemaps (including *titlebar* treemaps [16]) for topology depiction. Targeted tasks were categorized according to Cantu et al. [9] and the corresponding visual variables or attributes. The most common task types were *access information* tasks included in more than half of the studies, followed by *sort*, *distinguish*, and *compare* tasks. Least common task types were *associate*, *correlate*, and *identify distribution* tasks. Size-related tasks—part of more than half of the studies—represent the majority of used attributes, followed by tasks related to structure and color. With 45 %, the most common tasks were locate-a-node tasks and sort-by-size tasks with more than one third of the studies.

3.2 Summary of Findings

The results of the studies are thematically widely scattered and range from VTC to eye-movement exploration patterns. To extract and summarize these results, statistically significant findings of the studies were reviewed and verified instead of the respective claims. These summarized findings are not intended for citation as they are broadly reduced statements with a too general meaning—without further revalidation and replication, we insinuate that the measured effects are not generalizable by the studies' designs; refer to the original publications instead. The dependent variables measures are stated as follows: effectiveness (E), efficiency (I), and user satisfaction (S) [30].

Visualization Technique Comparison Studies. In comparison to recent visualization techniques, treemap performance results are heterogeneous. *Sunburst views* [4, 46] and *icicle plots* [4] perform better^{EIS} than treemaps for tasks that refer to tree attributes. Likewise, *icicle plots* perform better^{E} than treemaps for tasks related to counting and size [39, 40]. Titlebar treemaps and cushion treemaps perform better^{EIS} than *3D beamtrees* and *botanical trees* [26]—except for tasks related to level counting where 3D beamtrees perform better^{EI} [26, 50]. Findings targeting treemaps and *node-link diagrams* are diversified and partly indecisive. Traditional data visualization techniques were in most cases worse than treemaps when used on non-tree-structured data. Treemaps are more efficient than *tables* [16, 24] and have a higher accuracy than *pie charts* [21] for comparison tasks. For dense non-hierarchical sorted numerical data [60], treemaps perform worse^{E}. On the other hand, treemaps perform better^{E} than *bar charts* [27] for both high-density two-level hierarchies and comparison of non-leaf nodes. For search results displayed in a *list*, an additional treemap visualization improves the user's satisfaction and understanding of the results [12, 13].

Treemaps outperform other state-of-the-art exploration tools in terms of efficiency. Cushion treemaps, for example, are rated lower than *Windows Explorer* [1, 26] and perform worse^{EI} for tasks that refer to the complete (sub-)tree [26]. Then again, treemaps are more efficient and are rated better for these tasks than the *UNIX shell* [24,

Table 3: Visualization techniques that were compared to treemaps in VTC studies. Notation: evaluation with (★) and without (◇) significant results.
Abbreviation: State of the Practice (StotP).

	traditional				recent						StotP			
	List	Table	Bar Chart	Pie Chart	Nodelink	Sunburst	Icicle Plot	Beamtree	Botanical Tree	Windows Expl.	UNIX	Eclipse & Excel	Printout	Other
Directory Browsing [24, 49]											★			
GEnie Experiment [49]													★	
US Budget Experiment [24]		★												
3D Information Visualizations [58]					★									
Sunburst vs. Treemap (small) [46]						★								
Sunburst vs. Treemap (large) [46]						★								
Topology & Node Size [4]					★	★	★							
Compare to Beamtrees [50]								★						
Compare 5 Tree Visualizations [26]						★		★	★	★				
Treemaps & Tables [16]		★												
Crop Circles [53]						★								
Knowledge Discovery [54]						◇	★							
Hierarchy Browser [1]						★	★							
Search Engine Results [12]	★													
Multiple Views [47]	★													
Search Result Interface 1 [13]	★													
Search Result Interface 2 [13]	★													
Data Density [27]		★												
Node Link Treemaps [35]						◇			◇	◇				
Software Systems as Cities [57]												★		
Eye-tracking for EER-Tutor [36]						◇								◇
Rapid Overview [39]					★		★							
Software Landscape [15]					★									
Treemap vs. Euler [43]														◇
Treemap vs. Euler 2 [44]														◇
Photo: Layout [18]	★													
Compare 3 Tree Visualizations [40]						★	★							
Wrapped Bars: Comparison [60]			★											
Wrapped Bars: Ranking [60]			★											
Wrapped Bars: Distribution [60]			★											
Pie Charts: Proportion [21]				★										
Pie Charts: Comparison [21]				★										
Pie Charts: Labeling [21]				★										
Pie Charts & Stacked Bars [21]				★										

49]. Nested treemaps are more efficient for structure-similarity tasks than Windows Explorer [54]. For program comprehension tasks, treemaps perform better^{EI} than *Eclipse & Excel*, while being especially more efficient for leaf-node extrema tasks [57]. Likewise, treemaps are more efficient for global-scope tasks than *printout* financial reports [49].

Design Element Studies. Multiple user studies examined new and existing *visual variables* in the context of treemaps. *Luminance* was found to be suitable to encode secondary attributes when size encodes primary values [27] and for *color* the C-VPA coloring method showed improved size discrimination and higher aesthetic rankings than random coloring method [59]. Newly proposed visual variables are *sketchiness* in 2.5D [32] and *rectangular orientation* in 2D [31], both can be used as independent visual variables without noticeable impact on other visual variables. It should be noted that the accuracy of 2D *size* judgement tasks suffers when the compared rectangles have extreme *aspect ratio*s or when the rectangles are squares [27]. The findings of studies that compare *layout algorithms* are so extensive that it is difficult to summarize them in this report and are therefore future work [5, 19, 61].

The *depiction of topology* influences the perception of node size and structure [24]. For example, titlebar treemaps perform better^{EI} than cushion treemaps which perform better than non-nested
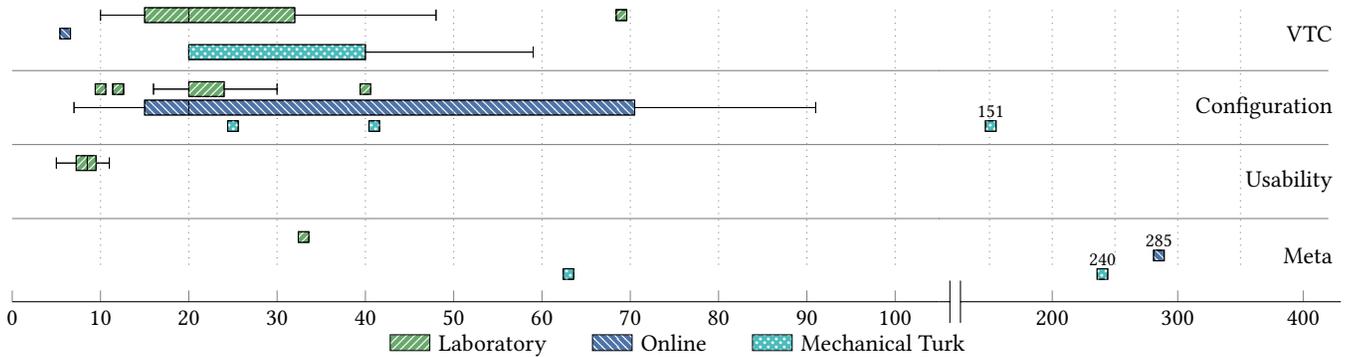
Figure 2: Number of participants categorized by evaluation practice and study implementation.

treemaps for tasks that refer to structure [6, 22, 26]. Likewise, titlebar treemaps perform better[IS] than simple nested treemaps for identification tasks and comparison tasks [16]. In addition to the kind of topology depiction, the *shape* changes the perception of structure and interpretation of relations. For relationship identification tasks, rectangular treemaps are less effective but are more efficient than map-like treemaps [6]. However, circular treemaps might have a better performance for structure extrema tasks than rectangular treemaps [53]. For different *dimensionalities*—especially 2D and 2.5D—differences could not be confirmed [8].

Suitable *focus and context* techniques for treemaps are Balloon Focus [48] and Drill-down [45]. Users of animated treemaps will take more likely shortcuts and rate them higher than users of nonanimated treemaps [7]. As a supporting technique, node *aggregation* is recognized by users and tasks that refer to determining nodes of interest are rated easier than without aggregation [33]. Similarly, in 2.5D treemaps, interactive *height-based filtering* might increase the accuracy of basic height-related tasks with only a marginal interaction overhead [34]. In *image treemaps*, users prefer fewer images and the presence of labels [18], while two-level-clustering reduces image finding time [17].

Usability Studies. During the exploration of a data set in a treemap visualization, certain *visual exploration strategies* apply. Treemaps are effective in promoting task-relevant fixations [40] and improve the user's perception of some patterns [54]. To explore these patterns in the underlying data treemap users used node, node count, node type, node size, structure, and content information [54]. However, treemap users constantly have to reorient and remember which elements were already processed [40]. The simultaneous usage of complementary visualization techniques may increase the user's understanding of the visualized data [47]. By their way of visualizing structure, treemaps are more or less suited for certain *task types*. In contrast to size-related tasks where treemaps offer a good sense of this attribute even between nodes of different levels [47], tasks related to global structure and some detailed structure-related tasks—level and sibling tasks—seem to be difficult to solve with treemaps [26].

Remaining findings include *shape*, *dimension*, and *interaction*. In map-like treemaps, users have a strong location recall with a clear understanding of islands, whereas landmarks might hinder interpretation of relationship discontiguity [20]. While in *virtual*

reality the interaction with 2.5D treemaps can be implemented by rotation, translation, and selection gestures [14]; 3D treemaps are suitable for exploration and long-term learning tasks [58].

Meta Studies. The task solving performance of visualization users is influenced by the compatibility of visual and *verbal metaphor*. This complex effect is varied by the user's preference of metaphor, spatial ability, and the locus of control [62, 63]. Beyond that, the visualization technique with the highest *visual quality* seems to perform relatively high in effectiveness-related measures [10].

Unconsidered Studies. Some user studies do not meet the requirements to be considered for further analysis. Due to a lack of a clear description of the results, the user studies on Euler diagrams were not considered [43, 44]. Second, we did not consider a study targeting the combination of treemaps and node-link diagrams since the methodical approach is not appropriate [35]. The authors tried to extend a previous user study [26] to compare their results with the other evaluation results. Due to a lack of comparability to the former user study—the used data set was different and the statistical data was not available—and no other comparable visualization that were evaluated within the same user study, the results were not considered. Similarly, we had to discard a user study using eye tracking as the visualized amount of data differed between the compared visualizations [36]. Last, we do not consider a user study targeting photo arrangements using a treemap approach [18], as the complex interactions and learning effects were not discussed.

4 CRITICAL REVIEW

Following, the treemap user studies are reviewed according to their quality, meaningfulness, and the resulting usefulness for the research community. The objective is to determine whether the findings can be accepted as valid and, thus, considered as generalized statements. If not, the main problems and shortcomings encountered in the studies are briefly analyzed.

4.1 Reviewing the User Studies

In order to assess the quality of these user studies individually as well as collectively, we reviewed the comprehensibility, fairness, and completeness of particular studies and the topic coverage across all studies. On the basis of the research goal and the hypotheses

derived, we have to conclude that many studies lack this information. Half of the user studies do not state their hypotheses based on a precise research goal, although they are the basis of every study. Several texts fail to comprehensively describe the study design and present their variable design. Almost no study discusses the interfering factors and the concluding validity issues. In addition to these design issues, the variable conditions in controlled experiments are often not randomized. Other frequently missing information includes detailed descriptions of data sets, participants, and procedure as well as values regarding statistical significance. About a quarter of the studies do not report on their participants' background. Typically, participants of treemap studies have none of domain knowledge, previous knowledge of treemaps, or experience therewith. Only a small amount of studies invited domain experts. The most common participant type are students of computer science or human-computer interaction. The number of participants is often too small to provide the statistical power or to serve the number of independent variables, as reported in previous surveys. Likewise, claims—especially stated in abstracts and conclusions—go far beyond the results. Even for studies that are not controlled experiments, results are sometimes verbally stated as valid generalized statements based on a causation.

Although comprehensive evaluation intends multiple points in time for assessment, the most common point is at or near the end of the development process. At this point, many researchers compare their approach with previous ones and try to prove that their approach is equally good at least, which requires a controlled experiment. This reduces the versatility of the study types used and thus the scope of the evaluation findings. Typically, such studies also test features of the user interface that are often not evaluated separately from the underlying visualization technique. Often, user studies derive their usage scenarios and tasks from previous publications, although no theoretical background is available. Since there is no standardized data set, each study uses own data set. Despite three consecutive studies by Ziemkiewicz et al. [62–64], there has not been one dedicated replication study on treemaps so far.

Task Compatibility. With the demand for fair control conditions, the configuration of a treemap depends on the task [24]. With the exception of Johnson's doctoral thesis [24] and configuration studies, most publications do not discuss whether the chosen configuration is appropriate. For tasks that refer to structure, treemaps with a larger offset or titlebar treemaps are appropriate, whereas size-related tasks require no offset [24]. Treemaps with cushion topology depiction [51] offer a trade-off between both needs. About 40 % of reviewed user studies contain tasks that refer to structure or tasks that refer to tree attributes, which are based on structure. Most of these studies configured their treemaps accordingly with offset or cushion. Whereas for the most common task attribute *size*, only one third of the studies are properly configured. The frequently used counting tasks—which appears in almost one fifth of the studies—in general do not fit visualization techniques that do have no support for a user interface. According to Ziemkiewicz, the matching of visual and verbal metaphor is important [63]. Although treemaps visually rely on the containment metaphor, about half of the studies used the level metaphor in their task descriptions. Treemaps are able to simultaneously show multiple visual variables

and the structure of a data set. Nevertheless, only one third of the studies included tasks that refer to two visual variables. Only four studies contain tasks targeting three dimensions, and correlation studies are the least common task type.

Coverage of Evaluation Topics. In order to assess the user studies as an entirety, we use the proposed evaluation topics to rate coverage and identify underrepresented areas. With respect to *Use Cases*, mainly VTC studies are relevant. Eleven techniques were compared and the results refer to three tasks types suitable for treemaps. An additional topic was the simultaneous visualization of treemaps and other techniques. While studies related to use cases include multiple studies, there is only one topic covered for *Communication*. With regard to *Interaction* aspects, the reviewed user studies cover pattern matching and multiple focus-and-context techniques including filtering in 2.5D, node aggregation, and animation. The connection of visual quality and the encouragement of users was also evaluated. The entirety of user studies considers several *Configurations*. From our assessment, questions regarding dimensionality are still open. The studies offer recommendations for shape, suitable to two tasks types extended by few qualitative insights. For the selection of an appropriate topology depiction, there are three recommendations related to different task types. In total, six visual variables were examined, whereby two were new in the context of treemaps. Treemap layout studies require a systematic review including non-participant measures [52]. Additionally, *Perceptional* aspects covered are task relevant fixations, area perception, and problems in orientation. Moreover, treemaps seem to lack a natural understanding of their basic containment metaphor without adequate training, such as the size of inner nodes [40]. In summary, treemap user studies suggested different appropriateness of treemaps in comparison to other visualization techniques related to six task types. While visual variables are well covered, there is hardly any validation of data set topology. The topics *Perception* and *Communication* are underrepresented. For treemap configuration, studies offer some suggestions regarding topology depiction, while dimensionality remains an open question. Further, treemap user studies seem to cover multiple aspects of treemaps but are fairly unstructured. Multiple topics are underrepresented, while others convey little structure.

4.2 Deriving Generalized Statements

Based on our review, we conclude that there is little that can be clearly derived. Various shortcomings and quality issues in study design, preparation, and execution were found. Other points to be added are the conclusiveness of research objective and study design as well as the analysis of nuisance or confounded variables, multiple participant related issues referring to count and representativeness of sample, and the lack of formal documentation of setup and conduct. In addition, almost no study addresses these problems in a validity discussion. There are also no replication studies. Concluding, this hinders general statements about treemaps, since most of the studies are affected. The presented study results should be used as hints and guides for future research and motivate replication studies of these. Researchers should thoroughly review previous studies before referring to their results or claims.

There is still little knowledge on which task types—and thus use cases or usage scenarios—are truly suitable to treemaps. Despite most-basic, theoretical background, there are no recommendations or guidelines on configuration that visualization designer could access. Likewise, many perceptional aspects of treemaps are unclear. For example, there is no resilient data on the value of simultaneous usage of more than one visual variable—although it is frequently referred to. Advantages or disadvantages of 2.5D treemaps or 3D treemaps are still unclear. Further, open questions are suitable ranges of data set sizes, learnability of the containment metaphor, and the mental-map characteristic of treemaps referring to long-term understanding of data. With regards to a holistic view on treemaps, the previous studies are historically grown and lack an inherent structure. Up to date we know of no research agenda to structurally prove previous statements about treemaps and evaluate the remaining aspects. This might be caused by missing references for a visualization evaluation standard, although there are initial thoughts on that topic [65].

5 CONCLUSION

Although a large number of user studies on treemap visualization techniques were designed, conducted, and published, there is still little that can be reliably derived or generalized. We addressed this gap by means of an analysis and review of the current state of research as well as a prospect for future researchers. We provided a summary, excerpt and review of 69 user studies that are additionally explorable on a community website. Although this list of studies is far from being complete, these resources can be used as a starting point for researchers to ease design, execution, and evaluation of user studies that target treemaps. We believe this evaluation, the review of previously published user studies, and the summary of guidelines serve as a starting point for future user studies and structure the evaluation process, leading to further insights into treemap visualization techniques.

As part of future work, we plan to prepare an overview of guidelines and conceptual models on user studies in information visualization to provide a systematic approach for researchers. A next step would be a research agenda with separate directions and projected effects. The provided data set is intended as a starting point and is target to get extended by the authors themselves to include evaluations of recently and future conducted user studies. Likewise, the insights from user studies targeting other tree and hierarchy visualization techniques are another area of interest for integration into the data set. But most importantly, we invite the information visualization community to collaborate on the user study data set.

ACKNOWLEDGMENTS

REFERENCES

[1] Keith Andrews and J. Kasanicka. 2007. A Comparative Study of Four Hierarchy Browsers using the Hierarchical Visualisation Testing Environment (HVTE). In *Proc. Conference on Information Visualisation (iV '07)*. IEEE, 81–86. https://doi.org/10.1109/IV.2007.8

[2] David Auber, C. Huet, A. Lambert, B. Renoust, A. Sallaberry, and A. Saulnier. 2013. GosperMap: Using a Gosper Curve for Laying Out Hierarchical Data. *TVCG* 19, 11 (2013), 1820–1832. https://doi.org/10.1109/TVCG.2013.91 IEEE.

[3] M. Balzer and Oliver Deussen. 2005. Voronoi treemaps. In *Proc. Symposium on Information Visualization (INFOVIS '05)*. IEEE, 49–56. https://doi.org/10.1109/INFVIS.2005.1532128

[4] T. Barlow and P. Neville. 2001. A comparison of 2-D visualizations of hierarchies. In *Proc. Symposium on Information Visualization (INFOVIS '01)*. IEEE, 131–138. https://doi.org/10.1109/INFVIS.2001.963290

[5] Benjamin B. Bederson, Ben Shneiderman, and Martin Wattenberg. 2002. Ordered and Quantum Treemaps: Making Effective Use of 2D Space to Display Hierarchies. *Transactions on Graphics* 21, 4 (2002), 833–854. https://doi.org/10.1145/571647.571649 ACM.

[6] Robert P. Biuk-Aghai, Patrick Cheong-Iao Pang, and Bin Pang. 2017. Map-like Visualisations vs. Treemaps: An Experimental Comparison. In *Proc. Symposium on Visual Information Communication and Interaction (VINCI '17)*. ACM, 113–120. https://doi.org/10.1145/3105971.3105976

[7] T. Bladh, D. A. Carr, and M. Kljun. 2005. The effect of animated transitions on user navigation in 3D tree-maps. In *Proc. Conference on Information Visualisation (iV'05)*. IEEE, 297–305. https://doi.org/10.1109/IV.2005.122

[8] Thomas Bladh, David A Carr, and Jeremiah Scholl. 2004. Extending tree-maps to three dimensions: A comparative study. In *Proc. Asia-Pacific Conference on Computer Human Interaction (APCHI '04)*. Springer, 50–59. https://doi.org/10.1007/978-3-540-27795-8_6

[9] A. Cantu, O. Grisvard, T. Duval, and G. Coppin. 2017. Identifying the Relationships Between the Visualization Context and Representation Components to Enable Recommendations for Designing New Visualizations. In *Proc. Conference on Information Visualisation (iV '17)*. IEEE, 20–28. https://doi.org/10.1109/iV.2017.55

[10] N. Cawthon and A. V. Moere. 2007. The Effect of Aesthetic on the Usability of Data Visualization. In *Proc. Conference on Information Visualisation (iV '07)*. IEEE, 637–648. https://doi.org/10.1109/IV.2007.147

[11] G. Chintalapani, Catherine Plaisant, and Ben Shneiderman. 2004. Extending the utility of treemaps with flexible hierarchy. In *Proc. Conference on Information Visualisation (iV '04)*. IEEE, 335–344. https://doi.org/10.1109/IV.2004.1320166

[12] Shixian Chu, Jinfeng Chen, Zonghuan Wu, Chee-Hung Henry Chu, and Vijay Raghavan. 2004. A Treemap-Based Result Interface for Search Engine Users. In *Proc. Symposium on Human Interface and the Management of Information*. Springer, 401–410. https://doi.org/10.1007/978-3-540-73345-4_46

[13] E. Clarkson, J. Foley, and K. Desai. 2009. ResultMaps: Visualization for Search Interfaces. *TVCG* 15, 6 (2009), 1057–1064. https://doi.org/10.1109/TVCG.2009.176 IEEE.

[14] Florian Fittkau, A. Krause, and W. Hasselbring. 2015. Exploring software cities in virtual reality. In *Proc. Working Conference on Software Visualization (VISSOFT '15)*. IEEE, 130–134. https://doi.org/10.1109/VISSOFT.2015.7332423

[15] Florian Fittkau, A. Krause, and W. Hasselbring. 2015. Hierarchical software landscape visualization for system comprehension: A controlled experiment. In *Proc. Working Conference on Software Visualization (VISSOFT '15)*. IEEE, 36–45. https://doi.org/10.1109/VISSOFT.2015.7332413

[16] Joseph H. Goldberg and Jonathan I. Helfman. 2005. Enterprise Network Monitoring Using Treemaps. *Human Factors and Ergonomics Society Annual Meeting* 49, 5 (2005), 671–675. https://doi.org/10.1177/154193120504900508 SagePub.

[17] A. Gomi, R. Miyazaki, Takayuki Itoh, and J. Li. 2008. CAT: A Hierarchical Image Browser Using a Rectangle Packing Technique. In *Proc. Conference on Information Visualisation (iV '08)*. IEEE, 82–87. https://doi.org/10.1109/IV.2008.8

[18] John Alexis Guerra-Gómez, Cati Boulanger, Sanjay Kairam, and David A. Shamma. 2016. Identifying Best Practices for Visualizing Photo Statistics and Galleries Using Treemaps. In *Proc. Working Conference on Advanced Visual Interfaces (AVI '16)*. ACM, 60–63. https://doi.org/10.1145/2909132.2909280

[19] Sebastian Hahn, Joseph Bethge, and Jürgen Döllner. 2017. Relative Direction Change – A Topology-based Metric for Layout Stability in Treemaps. In *Proc. Conference on Information Visualization Theory and Applications (IVAPP '17)*. SciTePress, 88–95. https://doi.org/10.5220/0006117500880095

[20] N. Hawes, S. Marshall, and Craig Anslow. 2015. CodeSurveyor: Mapping large-scale software to aid in code comprehension. In *Proc. Working Conference on Software Visualization*. IEEE, 96–105. https://doi.org/10.1109/VISSOFT.2015.7332419

[21] Lynn Huestegge and Tristan Herbert Pötzsch. 2018. Integration processes during frequency graph comprehension: Performance and eye movements while processing tree maps versus pie charts. *Applied Cognitive Psychology* 32, 2 (2018), 200–216. https://doi.org/10.1002/acp.3396 Wiley.

[22] Pourang Irani, Dean Slonowsky, and Peer Shajahan. 2006. Human Perception of Structure in Shaded Space-Filling Visualizations. *Information Visualization* 5, 1

(2006), 47–61. https://doi.org/10.1057/palgrave.ivs.9500113 Palgrave.

[23] T. Isenberg, P. Isenberg, J. Chen, M. Sedlmair, and T. Möller. 2013. A Systematic Review on the Practice of Evaluating Visualization. *TVCG* 19, 12 (2013), 2818–2827. https://doi.org/10.1109/TVCG.2013.126 IEEE.

[24] Brian Scott Johnson. 1993. *Treemaps: Visualizing Hierarchical and Categorical Data.* Ph.D. Dissertation. University of Maryland. UMI Order No. GAX94-25057.

[25] Brian Scott Johnson and Ben Shneiderman. 1991. Tree-Maps: A Space-filling Approach to the Visualization of Hierarchical Information Structures. In *Proc. Conference on Visualization (VIS '91)*. IEEE, 284–291. https://doi.org/10.1109/VISUAL.1991.175815

[26] A. Kobsa. 2004. User Experiments with Tree Visualization Systems. In *Proc. Symposium on Information Visualization (INFOVIS '04)*. IEEE, 9–16. https://doi.org/10.1109/INFVIS.2004.70

[27] N. Kong, Jeffrey Heer, and M. Agrawala. 2010. Perceptual Guidelines for Creating Rectangular Treemaps. *TVCG* 16, 6 (2010), 990–998. https://doi.org/10.1109/TVCG.2010.186 IEEE.

[28] R. Kosara, C. G. Healey, V. Interrante, C. Ware, and D. H. Laidlaw. 2003. User Studies: Why, How, and When? *Computer Graphics & Applications* 23, 4 (2003), 20–25. https://doi.org/10.1109/MCG.2003.1210860 IEEE.

[29] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. 2012. Empirical Studies in Information Visualization: Seven Scenarios. *TVCG* 18, 9 (2012), 1520–1536. https://doi.org/10.1109/TVCG.2011.279 IEEE.

[30] Gondy Leroy. 2011. *Designing User Studies in Informatics.* Springer. https://doi.org/10.1007/978-0-85729-622-1

[31] J. Liang, J. Hua, M. L. Huang, Q. V. Nguyen, and S. Simoff. 2012. Rectangle Orientation in Area Judgment Task for Treemap Design. In *Proc. Australian Conference on Computer-Human Interaction (OzCHI '12)*. ACM, 349–352. https://doi.org/10.1145/2414536.2414592

[32] Daniel Limberger, Carolin Fiedler, Sebastian Hahn, Matthias Trapp, and Jürgen Döllner. 2016. Evaluation of Sketchiness as a Visual Variable for 2.5 D Treemaps. In *Proc. Conference on Information Visualisation (iV '16)*. IEEE, 183–189. https://doi.org/10.1109/IV.2016.61

[33] Daniel Limberger, Willy Scheibel, Sebastian Hahn, and Jürgen Döllner. 2017. Reducing Visual Complexity in Software Maps using Importance-based Aggregation of Nodes. In *Proc. Conference on Information Visualization Theory and Applications (IVAPP '17)*. SciTePress, 176–185. https://doi.org/10.5220/0006267501760185

[34] Daniel Limberger, Matthias Trapp, and Jürgen Döllner. 2018. Interactive, Height-Based Filtering in 2.5D Treemaps. In *Proc. Symposium on Visual Information Communication and Interaction (VINCI '18)*. ACM, 49–55. https://doi.org/10.1145/3231622.3231638

[35] Lars Linsen and Sabine Behrendt. 2011. Linked treemap: a 3D treemap-nodelink layout for visualizing hierarchical structures. *Computational Statistics* 26, 4 (2011), 679–697. https://doi.org/10.1007/s00180-011-0272-2 Springer.

[36] Moffat Mathews, Antonija Mitrovic, Bin Lin, Jay Holland, and Neville Churcher. 2012. Do Your Eyes Give It Away? Using Eye Tracking Data to Understand Students' Attitudes towards Open Student Model Representations. In *Proc. Conference on Intelligent Tutoring Systems (ITS '12)*. Springer, 422–427. https://doi.org/10.1007/978-3-642-30950-2_54

[37] Liam McNabb and Robert S. Laramee. 2017. Survey of Surveys (SoS) - Mapping The Landscape of Survey Papers in Information Visualization. *Computer Graphics Forum* 36, 3 (2017), 589–617. https://doi.org/10.1111/cgf.13212 EG.

[38] L. Merino, M. Ghafari, C. Anslow, and O. Nierstrasz. 2018. A systematic literature review of software visualization evaluation. *Journal of Systems and Software* 144 (2018), 165–180. https://doi.org/10.1016/j.jss.2018.06.027 Elsevier.

[39] N. H. Müller, B. Liebold, D. Pietschmann, P. Ohler, and P. Rosenthal. 2013. Gaze into Hierarchy: A Practice-oriented Eye Tracking Study. In *Proc. EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization (EuroRVVV '13)*. EG, 9–10. https://doi.org/10.2312/PE.EuroRVVV.EuroRVVV13.009-010

[40] Nicholas Hugo Müller, Benny Liebold, Daniel Pietschmann, Peter Ohler, and Paul Rosenthal. 2017. Hierarchy Visualization Designs and their Impact on Perception and Problem Solving Strategies. In *Proc. Conference on Advances in Computer-Human Interactions (ACHI '17)*. IARIA, 93–101.

[41] Tamara Munzner. 2008. *Process and Pitfalls in Writing Information Visualization Research Papers.* Springer, 134–153. https://doi.org/10.1007/978-3-540-70956-5_6

[42] Catherine Plaisant. 2004. The Challenge of Information Visualization Evaluation. In *Proc. Working Conference on Advanced Visual Interfaces*. ACM, 109–116. https://doi.org/10.1145/989863.989880

[43] M. Sathiyanarayanan and N. Burlutskiy. 2015. Design and evaluation of euler diagram and treemap for social network visualisation. In *Proc. Conference on Communication Systems and Networks (COMSNETS '15)*. IEEE, 1–6. https://doi.org/10.1109/COMSNETS.2015.7098715

[44] Mithileysh Sathiyanarayanan and Nikolay Burlutskiy. 2015. Visualizing Social Networks Using a Treemap Overlaid with a Graph. *Procedia Computer Science* 58 (2015), 113–120. https://doi.org/10.1016/j.procs.2015.08.037 Elsevier.

[45] Kang Shi, Pourang Irani, and B. Li. 2005. An evaluation of content browsing techniques for hierarchical space-filling visualizations. In *Proc. Symposium on Information Visualization (INFOVIS '05)*. IEEE, 81–88. https://doi.org/10.1109/INFVIS.2005.1532132

[46] John Stasko, Richard Catrambone, Mark Guzdial, and Kevin McDonald. 2000. *An evaluation of space-filling information visualizations for depicting hierarchical structures.* Technical Report 5. Georgia Institute of Technology, Atlanta, GA, USA. 663–694 pages. https://doi.org/10.1006/ijhc.2000.0420 GVU Technical Report;GIT-GVU-00-03.

[47] Soon Tee Teoh. 2007. A study on multiple views for tree visualization. In *Proc. Conference on Visualization and Data Analysis (VDA '07, Vol. 6495)*. Society of Photo-Optical Instrumentation Engineers, 99–110. https://doi.org/10.1117/12.703076

[48] Y. Tu and H. W. Shen. 2008. Balloon Focus: a Seamless Multi-Focus+Context Method for Treemaps. *TVCG* 14, 6 (2008), 1157–1164. https://doi.org/10.1109/TVCG.2008.114 IEEE.

[49] David Turo and Brian Scott Johnson. 1992. Improving the Visualization of Hierarchies with Treemaps: Design Issues and Experimentation. In *Proc. Conference on Visualization (VIS '92)*. IEEE, 124–131. https://doi.org/10.1109/VISUAL.1992.235217

[50] Frank van Ham and Jarke J. van Wijk. 2003. Beamtrees: Compact Visualization of Large Hierarchies. *Information Visualization* 2, 1 (2003), 31–39. https://doi.org/10.1057/palgrave.ivs.9500036 Palgrave.

[51] Jarke J. van Wijk and Huub van de Wetering. 1999. Cushion treemaps: Visualization of hierarchical information. In *Proc. Symposium on Information Visualization (INFOVIS '99)*. IEEE, 73–78. https://doi.org/10.1109/INFVIS.1999.801860

[52] E. Vernier, M. Sondag, J. Comba, B. Speckmann, A. Telea, and K. Verbeek. 2020. Quantitative Comparison of Time-Dependent Treemaps. *Computer Graphics Forum* 39, 3 (2020), 393–404. https://doi.org/10.1111/cgf.13989

[53] Taowei David Wang and Bijan Parsia. 2006. CropCircles: Topology Sensitive Visualization of OWL Class Hierarchies. In *Proc. Semantic Web Conference (ISWC '06)*. Springer, 695–708. https://doi.org/10.1007/11926078_50

[54] Yue Wang, Soon Tee Teoh, and Kwan-Liu Ma. 2006. Evaluating the Effectiveness of Tree Visualization Systems for Knowledge Discovery. In *Proc. Symposium on Visualization (EUROVIS '06)*. EG, 67–74. https://doi.org/10.2312/VisSym/EuroVis06/067-074

[55] Colin Ware. 2013. *Information Visualization – Perception for Design.* Elsevier.

[56] Martin Wattenberg. 2005. A note on space-filling visualizations and space-filling curves. In *Proc. Symposium on Information Visualization (INFOVIS '05)*. IEEE, 181–186. https://doi.org/10.1109/INFVIS.2005.1532145

[57] Richard Wettel, Michele Lanza, and Romain Robbes. 2011. Software Systems As Cities: A Controlled Experiment. In *Proc. Conference on Software Engineering (ICSE '11)*. ACM, 551–560. https://doi.org/10.1145/1985793.1985868

[58] Ulrika Wiss and David A. Carr. 1999. An Empirical Study of Task Support in 3D Information Visualizations. In *Proc. Conference on Information Visualisation (IV '99)*. IEEE, 392. https://doi.org/10.1109/IV.1999.781587

[59] Yingtao Xie, Tao Lin, Rui Chen, and Zhi Chen. 2018. Toward improved aesthetics and data discrimination for treemaps via color schemes. *Color Research & Application* 43, 3 (2018), 328–340. https://doi.org/10.1002/col.22196 Wiley.

[60] Mehmet Adil Yalçın, Niklas Elmqvist, and Benjamin B. Bederson. 2017. Raising the Bars: Evaluating Treemaps vs. Wrapped Bars for Dense Visualization of Sorted Numeric Data. In *Proc. Graphics Interface Conference (GI '17)*. Canadian Human-Computer Communications Society, 41–49. https://doi.org/10.20380/GI2017.06

[61] Mengjie Zhou, Yan Cheng, Ning Ye, and Jing Tian. 2017. Effectiveness and Efficiency of Using Different Types of Rectangular Treemap as Diagrams in Cartography. In *Proc. Advances in Cartography and GIScience: Selections from the International Cartographic Conference.* Springer, 187–206. https://doi.org/10.1007/978-3-319-57336-6_14

[62] Caroline Ziemkiewicz, R. J. Crouser, A. R. Yauilla, S. L. Su, W. Ribarsky, and R. Chang. 2011. How locus of control influences compatibility with visualization style. In *Proc. Conference on Visual Analytics Science and Technology (VAST '11)*. IEEE, 81–90. https://doi.org/10.1109/VAST.2011.6102445

[63] Caroline Ziemkiewicz and Robert Kosara. 2008. The Shaping of Information by Visual Metaphors. *TVCG* 14, 6 (2008), 1269–1276. https://doi.org/10.1109/TVCG.2008.171 IEEE.

[64] Caroline Ziemkiewicz and Robert Kosara. 2009. Preconceptions and Individual Differences in Understanding Visual Metaphors. *Computer Graphics Forum* 28, 3 (2009), 911–918. https://doi.org/10.1111/j.1467-8659.2009.01442.x EG.

[65] Torre Zuk, Lothar Schlesier, Petra Neumann, Mark S. Hancock, and Sheelagh Carpendale. 2006. Heuristics for Information Visualization Evaluation. In *Proc. Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization (BELIV '06)*. ACM, 1–6. https://doi.org/10.1145/1168149.1168162