The Shaping of Information by Visual Metaphors

Caroline Ziemkiewicz and Robert Kosara

Abstract—The nature of an information visualization can be considered to lie in the visual metaphors it uses to structure information. The process of understanding a visualization therefore involves an interaction between these external visual metaphors and the user's internal knowledge representations. To investigate this claim, we conducted an experiment to test the effects of visual metaphor and verbal metaphor on the understanding of tree visualizations. Participants answered simple data comprehension questions while viewing either a treemap or a node-link diagram. Questions were worded to reflect a verbal metaphor that was either compatible or incompatible with the visualization a participant was using. The results (based on correctness and response time) suggest that the visual metaphor indeed affects how a user derives information from a visualization. Additionally, we found that the degree to which a user is affected by the metaphor is strongly correlated with the user's ability to answer task questions correctly. These findings are a first step towards illuminating how visual metaphors shape user understanding and have significant implications for the evaluation, application, and theory of visualization.

Index Terms—Cognition, visualization theory, metaphors, hierarchies, evaluation.

1 Introduction

Different information visualizations may present the same data in vastly different forms, as in the great variety of tree and graph visualizations. While these methods are often capable of showing the same information, it is widely recognized that any given method is better for some applications and worse for others. In cases where visualizations present equivalent information, it is the structural differences between methods that give rise to these differences in how the information they present can be used.

Understanding visualization, therefore, requires understanding how visualizations shape information, not only what information they present. A potential framework for this understanding is the formulation of a visualization as a set of visual metaphors. Metaphors are commonly used as a way of understanding how subtle differences in the form of language can suggest different interpretations of the same information [12]. By extending this idea to the interpretation of visual information, we can start building a framework for understanding how visual metaphors shape information (Figure 1).

To investigate the viability of this framework, we conducted an experiment to study the influence of metaphors on the processing of visual information. Our goal is to determine whether it makes sense to think of visualization in terms of visual metaphors, and if so, how these metaphors work to shape information. The results of our study indicate that priming a visualization user to think in an incompatible metaphor can slow her response time when responding to simple task questions. Furthermore, we found evidence that a user's accuracy in working with a visualization is related to her ability to internalize its visual metaphors. We then put these findings in context by considering the relevant research in diagrammatic reasoning and visual cognition in Section 7. This work serves as a foundation for a theory of visual information structure, which can enrich our understanding of visualization and its potential uses.

The contributions of this work include experimental evidence that metaphors affect the understanding and use of visualization, an informal meta-analysis of metaphorical influence on evaluations of tree visualization methods, and a survey that puts these findings in the context of cognitive research and graphical reasoning theory.

- *Caroline Ziemkiewicz is with UNC Charlotte, E-mail: caziemki@uncc.edu.*
- *Robert Kosara is with UNC Charlotte, E-mail: rkosara@uncc.edu.*

Manuscript received 31 March 2007; accepted 1 August 2007; posted online 27 October 2007.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

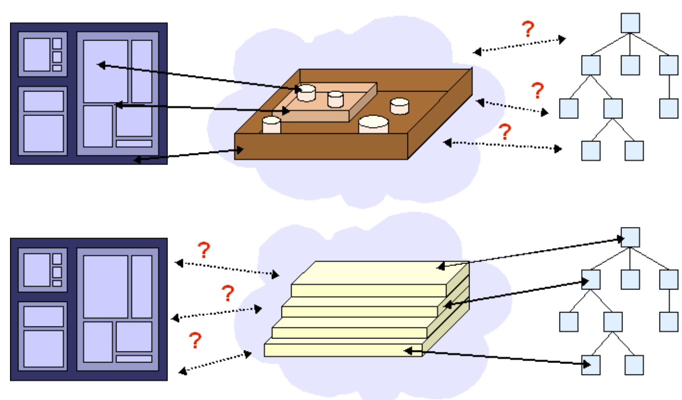


Fig. 1. Understanding a visualization involves matching its visual metaphor to internal metaphors. A treemap uses a metaphor of hierarchy as containment. If a user conceives of hierarchy data as a series of boxes nested inside each other, the correspondence to a treemap should be a natural fit, and the information will be easily understood. If she conceives of the hierarchy as a series of higher or lower levels, however, the correspondence will be more difficult and may take longer to process. For a visualization that employs a levels metaphor, like a node-link diagram, the reverse would be true.

2 Motivation

A better understanding of how visual structure shapes information can certainly contribute to the understanding of visualization as a field. It may also be an important factor in interpreting evaluation studies of visualization methods. As Chen's meta-analysis of visualization evaluation papers [5] suggests, there is little agreement in the findings of the various usability studies performed by visualization researchers over the years. This is no doubt due in large part to the lack of standard experimental procedures and benchmarks, but as Chen suggests elsewhere [4], another factor is the lack of understanding of the cognitive processes at work in visualization use. The traditional focus on simple information retrieval tasks obscures the complexity of human interaction with a visual representation.

Part of the inspiration for our current study lies in a potential consequence of this naïve view of evaluation. Hierarchy visualizations have been frequently subject to evaluation studies. Hierarchies are naturally applicable to a range of global and local information retrieval tasks, and many novel methods for visualizing them have been devised. However, there is little consensus across the existing evaluation papers in this domain, even when the same visualization methods and similar

Paper	Containment Questions	Treemap Ranking (Response Time)	Example Question (emphasis added)
Andrews and Kasanicka [1]	6 of 8 (75%)	1 out of 4	"Find the deepest subdirectory *inside* the directory 'pad++."'
Kobsa [11]	6 of 15 (40%)	2 out of 6	"Find the directory that *contains* the most .png type files."
Stasko et al. [17]	4 of 12 (33%)	2 out of 2	"Identify a directory *containing* files of a particular type."
van Ham and van Wijk [20]	0 of 5 (0%)	3 out of 4	"Users had to indicate *level* of a predetermined node."
Barlow and Neville [2]	0 of 5 (0%)	4 out of 4	"Participants counted the number of *levels* in the tree."

Table 1. Are variances in the metaphors of questions asked responsible for the inconsistent performance of a single visualization method across evaluation studies? We looked at several recent tree visualization evaluation papers that both included a treemap in their comparison and published their task questions. For each paper, we counted the number of questions or task descriptions that reflect a containment metaphor, i.e., those which used words like *contains*, *has*, *inside*, or *within* when describing relationships among nodes in the hierarchy. We then ranked each of the methods in the comparison by average response time overall, with 1 being the fastest method. Where exact response time was not reported, we estimated it based on results graphs. This preliminary meta-analysis suggests a possible relationship between metaphor compatibility and response time.

tasks are used. For example, in an evaluation by Kobsa [11], treemaps resulted in faster response times than any method besides Windows Explorer, while in a similar evaluation by Barlow and Neville [2], treemaps were the slowest method overall. More remarkably, the two studies included several highly similar task questions, leaving it nearly impossible to consistently interpret the treemap's strengths and weaknesses.

It is certainly possible that this inconsistency is entirely due to differences in the treemap implementation, the datasets used, the other methods studied, and the experimental designs. However, while the two studies used some logically equivalent retrieval tasks, we noticed a trend in the wording of the task questions. For example:

- Tree Balance
  - Kobsa: "Is the tree balanced or unbalanced? (A tree is unbalanced if its depth in one branch is at least two more than in another branch)."
  - Barlow and Neville: "Participants decided if the tree was balanced or unbalanced. Participants were told that balanced trees had leaves on the same level or two consecutive levels."
- Tree Depth
  - Kobsa: "What is the maximum depth of the eBay hierarchy?"
  - Barlow and Neville: "Participants counted the number of levels in the tree."

While the two pairs of questions are asking for the same information, Kobsa tended to word questions in terms of depth, and Barlow and Neville tended to word them in terms of levels. The latter study was the one which showed much worse performance for treemaps, particularly on these two questions. This suggests that priming participants to think about hierarchies in terms of levels may influence their ability to understand the treemap, which uses an unusual spatial metaphor of containment. A focus on levels would seem to favor a traditional node-link diagram or an icicle plot, which indeed were the two representations which performed best on these questions.

Based on the above, we conducted an informal meta-analysis of tree evaluation papers that published their task questions (see Table 1). The results reinforce our original hypothesis that consistent evaluation in information visualization is difficult because we lack a framework for including metaphorical structure in our interpretation of results. To further investigate this possibility, we designed an experiment to study the effect of metaphor on visualization use.

3 Hypothesis

The question of how a visual metaphor can be said to shape the information in a visualization is a broad and challenging one. Based on our meta-analysis, we chose to investigate two more specific questions:

1. Are visual metaphors analogous to verbal metaphors?
2. Does priming a user with a particular verbal metaphor affect her ability to process an analogous visual metaphor?

Since verbal metaphors are known to influence how information is processed, we can use verbal instructions to prime a user to think in terms of a particular metaphor. We can then test whether a visual metaphor influences thought in the same way by testing whether that priming affects the speed at which visual information is understood.

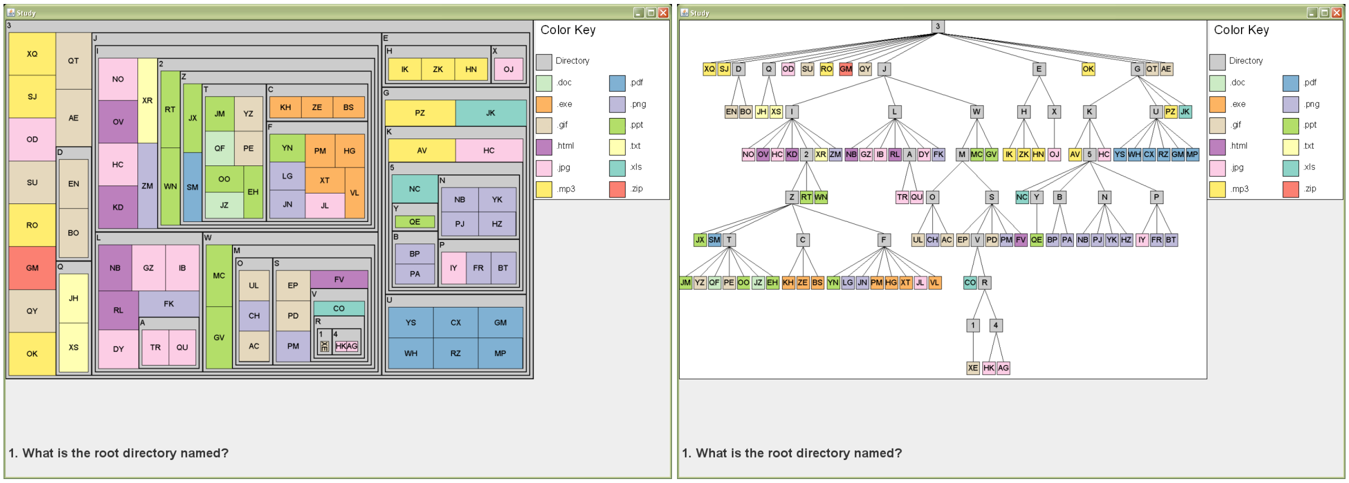
Another reason to consider the interaction of verbal and visual metaphor is a more pragmatic one. Approaches to understanding visualization without reference to how visual representations structure information are limited in their capability to incorporate the effects of context, prior knowledge. and other outside factors on the use of visual information. This may be a factor in the existing issues with evaluation of information visualization methods: we lack a framework to consider how a user's existing structures, and those that come from other parts of the environment, affect the knowledge we derive. The language we use to describe a problem is one of those outside factors, and one that may be affecting current attempts to evaluate information visualization methods.

Therefore, our hypothesis is that participants will be slower in responding to questions that reflect a verbal metaphor which is incompatible with the visual metaphor of the visualization they are using and faster in responding to questions that reflect a compatible verbal metaphor.

4 Methods

To study the effects of verbal metaphor and visual metaphor on task performance, we performed a study with 33 participants, all students at the University of North Carolina at Charlotte. Students were recruited from architecture and computer science courses, and included 9 females and 24 males, with age ranging from 18 to 40 (M = 23.4). Students were primarily undergraduates, and about half (51.5%) were in computing-related majors. All but one rated their comfort with computers as "fairly comfortable" or higher, and all reported using a computer at least once or twice a day. Only three participants said they had used an information visualization before.

During the study, participants were shown three hierarchical datasets in one of two tree visualizations: a treemap (Figure 2a) or a node-link diagram (Figure 2b). The visualization type varied between subjects, so that each participant saw only one type of visualization. The visualizations were created using the Infovis Toolkit [7], and we attempted to keep as many of the surface visual qualities constant



(a) Treemap condition (b) Node-link condition

Fig. 2. These screenshots show our study design as it was seen by participants. Each participant viewed one of two types of tree visualization. Using the InfoVis Toolkit, we kept superficial qualities of the visualizations (e.g., color, label size, and display size) as constant as possible. The visualizations were not interactive. Participants answered questions about the data by pressing keys on the keyboard; for each of three sessions, participants saw eight questions, four of which were metaphorically compatible with their visualization and four of which were incompatible. The order of compatible and incompatible questions was random. Each session showed a different but similarly complex dataset based on a hypothetical file structure.

across the two representations as possible. The color scheme, window size, and label appearance did not differ between the two. Additionally, node size in the treemap was not given a meaning, so that as much as possible, the two visualizations expressed the same information. In order to focus entirely on the effect of visual representation, we did not include interactivity in either visualization.

The three datasets were described to the participants as representing hypothetical file hierarchies. Color was used to indicate file type, and the same color key was provided to the two groups. Files were named with random strings of two characters, and directories were named with a single random character or digit. These random names were meant to remove any possibility of the user answering the questions through inference rather than by consulting the visualization, as may be possible when using meaningful file and directory names.

Participants were initially told that the purpose of the study was to evaluate different types of hierarchy visualization. After an initial training period in which they answered four questions and were given a chance to try again if they answered incorrectly, participants were asked eight questions about each dataset. Each question was answerable with a single keystroke: either a number or one of the single-character directory names. During the experimental phase, participants were not informed if they answered incorrectly.

For each of the eight questions, we prepared two versions: one that reflected a "containment" metaphor, and one that reflected a "levels" metaphor (Table 2). The containment metaphor was considered to be more compatible with the treemap view, and the levels metaphor was considered to be more compatible with the node-link view.

Verbal metaphor was varied within subjects, in order to study the compatibility effect independently of individual differences in accuracy and response time. During their time with each of the three datasets, a participant saw four questions of the containment type and four questions of the levels type. The set of questions used for each dataset was counterbalanced from subject to subject, and question order was randomized. The result is that each participant, during each session, would answer a series of eight questions that randomly varied between a compatible and an incompatible metaphor relative to the visualization she was using.

For each question, we measured the participant's response time and whether they answered the question correctly. Altogether, participants answered twenty-four task questions. After the three sessions were complete, users filled out a short usability survey and were asked to write any comments about the visualization they had used.

5 Results

The results of our study, measured in response time and accuracy, suggest a complex picture of how metaphors affect visualization use. Participants do show slower response time when responding to task questions with metaphors incompatible with the visualization they are using, an effect which held across both visualization types and both verbal metaphor types. More surprisingly, we found that the amount by which any given participant performed faster on compatible questions was strongly correlated with that participant's overall accuracy.

5.1 Metaphor Compatibility and Response Time

To test the hypothesis that participants will perform faster on questions compatible with their visualization than on those which are incompatible, we first computed the participants' overall mean response time on incompatible task questions and on compatible task questions, with only correct responses considered. When the means are compared across the three sessions of our experiment, there is a clear training effect. In the first two sessions, the compatible responses (M = 22.5s, SD = 13.9s) are indeed slightly faster than those for the incompatible responses (M = 23.5s, SD = 17.4s). However, not only do both the incompatible and compatible means show a significant decrease over time ($F(2, 194) = 3.79$, $p < 0.05$), but the difference between them also decreases; in fact, the incompatible mean decreases faster than the compatible one, so that the effect slightly reverses by the third session (Figure 3). Although there are trends in the means, neither of these effects reach significance due to the high degree of variance. In order to further examine the differences in responses independently of the training effect, our following statistical analyses only included responses from the first two sessions.

Because of the way we designed the experiment, the task questions that participants answered in Session 1 and Session 2 were exactly symmetrical in terms of metaphor; that is, if they received a question in Session 1 with a levels metaphor, they would receive the same question in Session 2 with a containment metaphor. This means that we can perform a repeated measures test in which we compare each participant's response time on a single question in both the compatible and the incompatible case. We performed a paired samples t-test

Containment Metaphor	Levels Metaphor
1. How many directories enclose the deepest file?	1. How many directories are above the lowest-level file?
2. How many total subdirectories are within the directory "S"?	2. How many total subdirectories are under the directory "S"?
3. How many files are immediately inside the directory "I"?	3.How many files are immediately under the directory "I"?
4. What is the deepest directory that contains both "XE.gif" and "KH.exe"?	4. What is the lowest-level directory that both "XE.gif" and "KH.exe" fall under?
5. Which directory immediately contains the most files of type ".pdf"?	5. Which directory can the most files of type ".pdf" be found immediately under?
6. What is the directory that immediately contains the directory "V"?	6. What is the directory immediately above the directory "V"?
7. Which directory contains the largest number of immediate subdirectories?	7. Which directory has the largest number of subdirectories immediately below it?
8. Which directory contains a deeper hierarchy: "G" or "M"?	8. Which directory has more levels under it: "G" or "M"?

Table 2. The eight task questions given to participants in our study, framed in either a metaphor of hierarchy as containment or hierarchy as levels, as asked during the first trial session. For the subsequent sessions, the specific files and directories mentioned in the questions were altered to match the dataset being visualized, but the wording remained unchanged. For each session, a participant saw four questions from the Levels list and four from the Containers list, and question order was randomized. These questions are based on common task questions asked in tree visualization evaluation papers.

that compared these pairs of responses and found a stronger trend, $t(139) = 1.05$, $p = 0.297$ (Figure 4). While this trend is not statistically significant, it may suggest that metaphor compatibility has some effect on users' speed of processing.

However, there is a large amount of individual variance in performance among the participants. Some participants' response times favored the compatible metaphors by a very large amount on average, and some mostly favored the incompatible metaphor. To our surprise, this difference is not predicted by the type of visualization they were using. There were also no correlations with the participant's gender or whether they were in a computing-related major. Other factors which were ruled out included differences in compatibility effect that might arise from the questions themselves. While some questions took longer to answer overall than others, there was no evidence that any question was more difficult in one metaphor or another. It is interesting that the more difficult questions seemed more likely to show a strong effect of metaphor compatibility (e.g., questions 4 and 7, in Figure 4).

5.2 Metaphor Use and Accuracy

One factor which did predict whether a given participant performed faster on metaphorically compatible questions, however, is overall accuracy. The average difference between a participant's incompatible response time and compatible response time highly correlates with that participant's number of correct responses across all three sessions, $r(31) = 0.49$, $p < 0.01$. That is, the degree to which a participant favors the compatible metaphor strongly correlates with that participant's accuracy in using the visualization (Figure 5). When participants are grouped based on the number of questions they answered correctly, it is clear that participants who answered fewer questions correctly are in fact more likely to respond faster to the incompatible questions (Figure 6).

Furthermore, when controlling for the number of correct responses using a repeated measures analysis of covariance (ANCOVA), we did find a significant effect of compatibility on a participant's response time, $F(1, 26) = 9.05$, $p < 0.01$. Taken together, these findings suggest a close relationship between a user's understanding of a visualization and her ability or inclination to internalize its visual metaphors.

6 Discussion

The findings of this study strongly support the need for better consideration and understanding of how structure influences the processing of information visualization. Our discovery of a training effect in terms of metaphorical influence has implications for how a visualization is learned and integrated into a user's thinking process. Additionally, the fact that metaphorical influence is correlated with user accuracy suggests that internalizing the visual metaphor is an important step in using any given visualization.

6.1 Structure Influences Understanding

The results of our paired-response test suggest that there can be significant interference between visual and verbal metaphors when understanding information, and that for some users, this interference can cause delays of several seconds when performing simple tasks. At the very least, this effect needs to be considered when designing future evaluations of visualization methods.

More broadly, this effect suggests that a translation must be taking place between the visual and verbal representations a given user holds about the information in a visualization, and that this translation can be influenced by the metaphors in play on both sides. This supports the importance of the conceptual structure of information we glean from a visualization; if this information were entirely uninfluenced by the visual representation it was drawn from, one would not expect to find any difference in performance between the incompatible and compatible cases. The fact that there may be a reliable difference means not only that the metaphorical structure of the visualization is recorded, but also that this structure influences how we are able to use the information we gain from a visualization.

It is also of interest that we found no differences in the amount of compatibility effect between users of the two visualization types, even though node-link diagrams are likely to be more familiar and treemaps are arguably more tied to an unusual visual metaphor. This suggests that the effect of metaphor is a constant factor across all types of visualization, even common ones.

6.2 Learning and Abstraction

The fact that the compatibility effect decreases over time while using a visualization is of interest. It is ambiguous from our study design whether this arises from the participants' growing familiarity with the visualization, the task questions, or the combination of the two. Curiously, the average response to incompatible questions was actually slightly *faster* in the third session (Figure 3). An intriguing possibility is that, over the course of viewing a single visual metaphor and being asked both compatible and incompatible questions about it, a user gets better at abstracting or translating between the metaphors. While we cannot draw any conclusions on this matter from the current study, this is a clear area for future research.

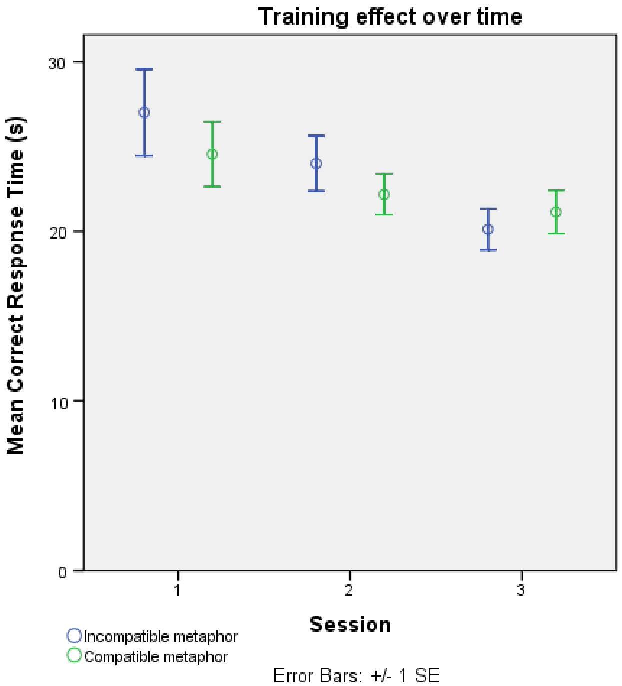


Fig. 3. The mean response times across all participants for both compatible and incompatible questions across the three sessions. Only correct response times are considered. While there is a difference in means between the compatible and incompatible conditions for the first two sessions, the overlapping error bars suggest that this isn't a statistically significant effect. The change in means over time shows that participants improved their response time over the course of the study, and also that the compatibility effect decreases with experience. This may suggest that participants get better at abstracting or translating metaphors while using visualizations.

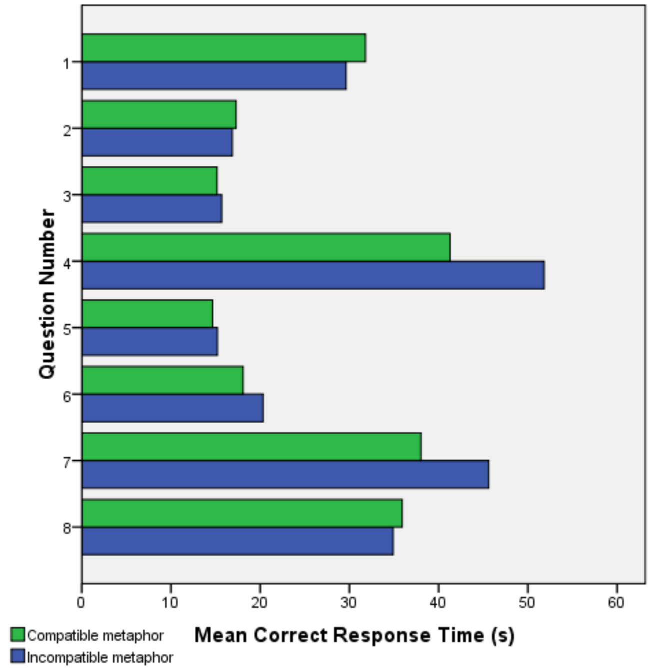


Fig. 4. This shows the paired response differences for each of the eight questions (Table 2) across the first two sessions. For each participant and each question, we take the participant's response time on an incompatible version of the question and on a compatible version of the question. These bars show the averages for all participants for each of the eight questions. Since both participant and question are held constant, this is the most fine-grained way to look at the effect of compatibility alone. While there is a great deal of variation across questions, the most dramatic differences show a slower response time for incompatible questions.

6.3 Internalizing the Metaphor

Perhaps the most compelling finding in our study is the correlation we discovered between a participant's metaphorical influence and accuracy. Whether or not visual metaphor can be said to reliably affect all users' conceptual knowledge structures, this effect suggests that users who do attempt to internalize the visual metaphor may be better able to use the visualization effectively. It is possible that those who performed faster on the incompatible questions were those who already had a strong preference for a metaphor which opposed the visualization they were using, and this is why they had difficulty answering questions with that visualization.

While we did not gather data on how each participant naturally conceives of hierarchies, some support for this speculation can be found in the comments made by users after completing the study. One treemap user with 10 out of 24 correct responses, an average response time of 15.3 seconds for incompatible questions, and an average response time of 23.5 seconds for compatible questions, wrote, "The lack of white space made seeing the different levels hard." Another had 7 correct responses and an average response time of 6.2 seconds for incompatible and 20 seconds for compatible questions, and mentioned a preference for flow charts for displaying data. If these participants were thinking in levels or flow metaphors while trying to use a treemap, it may have been harder for them to use the visualization overall, explaining their low accuracy. On the other hand, questions framed in a metaphor closer to their own conceptualization may have required one less step of translation, leading to faster response times in those cases.

This finding has strong implications for the design and application of visualization. Users may not only need to learn how to read a particular visualization in order to use it, but also to incorporate its particular metaphors into their own thinking. Certainly this would suggest that more immediately successful visualizations are likely to be those that match the user group's existing metaphors about their data and the work they need to perform with it, and that the discovery of these existing metaphors should be an important part of the visualization design process. That said, given the indications in this work that visual representation can work to restructure internal representation, there may be advantages to introducing novel metaphors that provide new perspectives on existing problems. Striking a balance between metaphorical familiarity and the benefits of reconceptualizing problems may prove to be an ongoing challenge in visualization practice.

7 Related Work

There exists an extensive, if sporadic, body of work on how the structure of diagrams and other visual representations shapes our understanding of their informational content. General research on how people reason using diagrams and other visual representations of information provides a theoretical context for our findings, and work on the interaction between visual and verbal representations helps to explain the interference effect we found in metaphor compatibility.

7.1 Evaluation in Information Visualization

blah blah

7.2 Diagrammatic Reasoning

An influential theory in this regard comes from Pinker, who considers the structure of a diagram in terms of its "graph schema," or the system of mappings between data dimension and visual dimension that define it [15]. Pinker's view, and others like it, tend to assume an information extraction perspective; information is latent in the diagram, and the

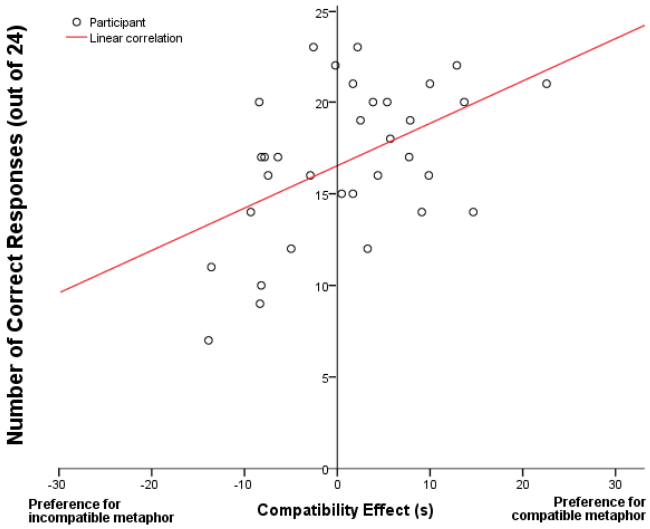


Fig. 5. We calculate each participants preference for compatible metaphors by subtracting their average response time on compatible questions from their average response time on incompatible questions. High positive values therefore indicate a strong preference for compatible metaphors, while negative values indicate a preference for incompatible metaphors. When participants are plotted in terms of their preference and overall accuracy, there is a strong correlation between preference for compatible metaphors and accuracy, $r(31) = 0.49$, $p < 0.01$.

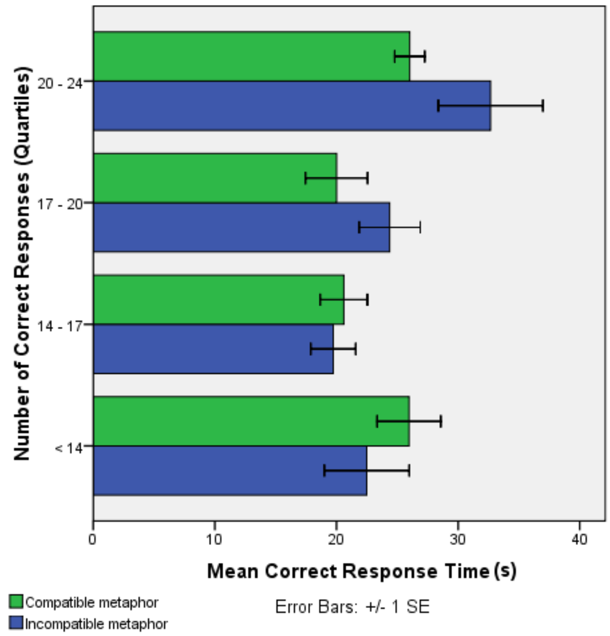


Fig. 6. Here we divide the participants into quartiles based on the number of questions they answered correctly (out of 24). Participants in higher-accuracy groups have a much higher tendency to perform faster on metaphorically compatible questions, while those in the lower accuracy groups tend to perform faster on incompatible questions. This suggests that an inability to internalize the visual metaphor of a visualization is linked to an inability to extract information from the visualization accurately.

process of understanding a visualization is translating that information out and reasoning with it in the head. There can be differences between two informationally equivalent visual representations in that one can make certain information more salient than the other. But because of their focus on one-way processing, these models do not readily allow for the influence of internal structures on how visualizations are understood.

Cleveland and McGill [6] exemplify the focus on information extraction as the primary process in understanding diagrams. Their work has served as the basis for guidelines on which low-level visual mappings are appropriate for which sort of data, but they do not take overall visual structure into account when analyzing differences in graphs.

Other theories focus more on reasoning with visual representations, such as Stenning and Oberlander's view of diagrams and language as logically equivalent yet supporting different facilities of inference [18]. Similarly, Larkin and Simon [13] consider the differences between graphical and verbal representations as differences of what information is made salient and explicit. The authors consider what effects the structure of a representation has on understanding, although they focus on the very broad differences between words and pictures rather than defining differences among types of graphical structure.

An important contribution in this regard has come from the extensive body of work by Tversky and colleagues, which includes numerous experiments on how people interpret information presented in different visual representations. For example, the authors presented the same simple two-point data as either a bar chart or a line graph and asked for users' interpretations [21]. They found that those viewing a bar chart tending to describe the diagram as depicting two separate groups, while those viewing a line graph described the data as a trend. This effect held even when the interpretations conflicted with the labels on the data points. These findings and others are further discussed as examples of how schematic figures such as bars and lines are interpreted in varying contexts [19].

This idea of context is very close to what we mean by metaphor, that is, a set of structural properties that provide a framework for meaning. As Tversky et al. write, context aids the interpretation of ambiguous primitive features such as blobs and lines by fitting their relevant properties to task demands. The function of metaphor is similar, but adds an extra layer of interpretation in the implied logical properties of the metaphor source [**?**]. This allows outside information and inferences to be applied to a visualization. In this sense, metaphors are very close to the idea of mental models in user interface and industrial design; as discussed by Kieras and Bovair [10], users with a meaningful model of how and why a device works are not only more able to remember memorized tasks using the device, but are also more likely to spontaneously find a more efficient way to perform the task. A mental model of this kind may not be applicable in the more open-ended task of understanding a dataset, but a strong metaphor for the data may aid understanding in the same way.

Gattis and Holyoak [8] discuss the ways in which diagrams function as external representations that mediate between internal representations of a visual and conceptual nature, in reference to a experiment showing that the mapping between spatial dimensions and meaning affects participants' accuracy in reading the graph. They found that the ideal layout in terms of increasing user accuracy was more affected by task demands than by any constraint that was constant across tasks. A similar finding is reported by Jarvenpaa in the context of decision-making based on graphical representations [9]. Jarvenpaa's work highlights the importance of context and other outside factors in visualization comprehension, and is an important precursor to our research.

7.3 Visual and Verbal Structure

Unlike work that investigates how varying task demands interact with varying visual representations, our findings refer to cases where task demands are constant but a user's internal knowledge representations vary. Questions about the nature of knowledge representation are lively and controversial ones in current cognitive science, and it is not within the scope of this work to tackle the issue. But if at some level, working with a visual representation entails a translation from visual representation to verbal representation and vice versa, understanding the interaction between those two modes may provide a first step towards understanding how visualization works to increase or alter knowledge.

Bryant considers the interaction between language and perception as they contribute to the understanding of space, and argues for a common spatial representation underlying both [3]. Experiments by Oakhill and Johnson-Laird [14] that show encoding verbal descriptions of spatial scenes is more difficult when a participant is simultaneously carrying out a visual tracking task support this claim. Likewise, Spivey et al. [16] recently found evidence for a strong and rapid influence of verbal instructions on perception of a visual scene. These findings suggest the importance of interaction between information gained visually and verbally, and support our decision to use metaphors as a framework for understanding the structure of visual representation.

8 Conclusions and Future Work

This work represents a significant first step towards an understanding of how the structure of a visualization influences how we process it. By showing that compatibility of visual metaphor and verbal metaphor can speed the processing of visual information, we provide evidence that visual metaphor influences the representation of information in the mind. Further, we have shown that the degree of this compatibility effect is closely associated with a user's ability to answer questions based on the visualization, which suggests that internalizing the visual metaphor is an important component of visualization comprehension.

However, the results of this study do not clarify what the underlying cause of this correlation might be. Further experiments are needed to isolate this effect and find its source, and to examine whether there are factors (such as spatial ability, visual literacy, or prior knowledge) that can predict whether a user successfully internalizes a visual metaphor. The fact that neither the traditional node-link diagram or the more novel treemap was more associated with the compatibility effect suggests that mere familiarity cannot be the only factor.

Another effect found in this study that bears further research is the suggestion that the difference between compatible and incompatible response times decreases over time, and that in some cases incompatible results have an advantage by the third session. Although we did not find a significant result in this study, a more focused experiment may be able to isolate and explain this finding. Further investigation of this effect may shed light on how a visualization is learned and how learning a visualization affects a user's understanding of information.

Our evidence points towards a theory of visualization as the construction of knowledge through the fitting of information to internal and external metaphors. But more work is needed to fully establish this theory. A major open question is how to define the metaphors inherent in an existing visualization. Seeing a treemap as a set of containers may be fairly intuitive, but what is the inherent metaphor in BeamTrees? Or in a radial tree layout? And what about visualizations that mix metaphors? One can certainly come up with possible answers to these questions, but validation is necessary if these metaphors are to be useful in practice. One possibility is to use a process similar to the procedure of this study to test hypotheses about metaphors: if the correlation is apparent, the metaphor fits. Further replication and validation of this effect may lead to a standard process for metaphor testing and improvement.

A limitation of the current experiment is that we chose to study visual representations without interactivity. While this helped to isolate the effects of visual representation as it relates to very low-level information retrieval tasks, a more complex picture will undoubtedly emerge if interaction is taken into account. Visual metaphors may suggest potential interactions to a user on their own, as is commonly exploited in user interface design. A related question is whether interacting with a visualization affects a user's sense of its structure and ability to abstract its metaphors. These questions are certainly deserving of further study.

Validation of the metaphorical effect, a better understanding of what causes it, and ways to describe and categorize visual metaphors will help lay the foundation for a theory of visualization based on the structuring of information through visual metaphors. Continued exploration of the implications and nuances of this theory will contribute to the understanding of how and why visualization works.

Acknowledgements

This work was performed with partial support from the National Visualization and Analytics Center (NVAC™), a U.S. Department of Homeland Security Program, under the auspices of the SouthEast Regional Visualization and Analytics Center. The authors wish to thank Paula Goolkasian for her helpful comments on study design, Eric Sauda and Amy Ulinski for their kind assistance with participant recruitment, Evan Suma for his suggestions on statistical procedures, and Remco Chang for many fruitful discussions.

References

[1] K. Andrews and J. Kasanicka. A comparative study of four hierarchy browsers using the hierarchical visualisation testing environment (hvte). In *Proceedings International Conference Information Visualization (IV)*, pages 81–86. IEEE CS Press, 2007.

[2] T. Barlow and P. Neville. A comparison of 2-d visualizations of hierarchies. In *Proceedings Information Visualization*, pages 131–138. IEEE CS Press, 2001.

[3] D. J. Bryant. Representing space in language and perception. *Mind & Language*, 12(3/4):239–264, 1997.

[4] C. Chen. Top 10 unsolved information visualization problems. *IEEE Computer Graphics and Applications*, 25(4):12–16, 2005.

[5] C. Chen and Y. Yu. Empirical studies of information visualization: A meta-analysis. *International Journal of Human-Computer Studies*, 53:851–866, 2000.

[6] W. S. Cleveland and R. McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387):531–554, 1984.

[7] J.-D. Fekete. The infovis toolkit. In *Proceedings Information Visualization*, pages 167–174. IEEE CS Press, 2004.

[8] M. Gattis and K. J. Holyoak. Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22(1):231–239, 1996.

[9] S. L. Jarvenpaa. The effect of task demands and graphical format on information processing strategies. *Management Science*, 35(3):285–303, 1989.

[10] D. E. Kieras and S. Bovair. The role of mental models in learning to operate a device. *Cognitive Science*, 8:255–273, 1984.

[11] A. Kobsa. User experiments with tree visualization systems. In *Proceedings Information Visualization*, pages 9–16. IEEE CS Press, 2004.

[12] G. Lakoff and M. Johnson. The metaphorical structure of the human conceptual system. *Cognitive Science*, 4:195–208, 1980.

[13] J. H. Larkin and H. A. Simon. Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11:65–99, 1987.

[14] J. V. Oakhill and P. N. Johnson-Laird. Representation of spatial descriptions in working memory. *Current Psychology*, 3(1):52–62, 1984.

[15] S. Pinker. A theory of graph comprehension. In R. O. Freedle, editor, *Artificial Intelligence and the Future of Testing*, pages 73–126. Lawrence Erlbaum Associates, 1990.

[16] M. J. Spivey, M. J. Tyler, K. M. Eberhard, and M. K. Tanenhaus. Linguistically mediated visual search. *Psychological Science*, 12(4):282–286, 2001.

[17] J. Stasko, R. Catrambone, M. Guzdial, and K. McDonald. An evaluation of space-filling information visualizations for depicting hierarchical structures. *International Journal of Human-Computer Studies*, 53:663–694, 2000.

[18] K. Stenning and J. Oberlander. A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, 19(1):97–140, 1995.

[19] B. Tversky, J. Zacks, P. U. Lee, and J. Heiser. Lines, blobs, crosses and arrows: Diagrammatic communication with schematic figures. In *Theory and Application of Diagrams*, pages 221–230. Springer-Verlag, 2000.

[20] F. van Ham and J. J. van Wijk. Beamtrees: Compact visualization of large hierarchies. In *Proceedings Information Visualization*, pages 93–100. IEEE CS Press, 2002.

[21] J. Zacks and B. Tversky. Bars and lines: A study of graphic communication. *Memory & Cognition*, 27(6):1073–1079, 1999.