SMITH COLLEGE

Smith ScholarWorks

11-15-2012

Understanding Visualization by Understanding Individual Users

Caroline Ziemkiewicz
Brown University

Alvitta Ottley
Tufts University

R. Jordan Crouser
Tufts University, jcrouser@smith.edu

Krysta Chauncey
Tufts University

Sara L. Su
Google Inc.

See next page for additional authors

Follow this and additional works at: https://scholarworks.smith.edu/csc_facpubs

Part of the Computer Sciences Commons

Authors

Caroline Ziemkiewicz, Alvitta Ottley, R. Jordan Crouser, Krysta Chauncey, Sara L. Su, and Remco Chang

Understanding Visualization by Understanding Individual Users

Caroline Ziemkiewicz[*]
Brown University

Alvitta Ottley[†]
Tufts University

R. Jordan Crouser[‡]
Tufts University

Krysta Chauncey [§]
Tufts University

Sara L. Su [¶]
Google Inc.

Remco Chang [‖]
Tufts University

◆

1 THE CHALLENGE OF INDIVIDUAL DIFFERENCES

Visualizations are tools to support thinking. They can be used to externalize knowledge about a complex analytical task or domain, and through interaction, they can embody a reasoning process. As such, visualization cannot be fully understood without also understanding how the user of a visualization thinks. This understanding is nontrivial, and has been complicated by mounting evidence that there is no single type of visualization user. Ultimately, making sense of visualization requires understanding how users vary and why.

Past research in visualization theory has focused primarily on how data can be mapped to visual forms and how people perceive them. These research endeavors have led to the identification of fundamental principles regarding how humans perceive colors and visual patterns, and have led to the establishment of general design guidelines for developing useful visualizations. Perceptual visualization theory attempts to understand and model how users perform fundamental low-level tasks.

However, as visualization gains widespread importance, the tasks that researchers must study are becoming more complex. In recent years, visualizations are being used as cognitive aids in problem solving, as users come to rely on visualizations to help them solve increasingly difficult problems. While color and perceptual theories remain necessary to make good design decisions, by themselves they are not sufficient to guide the design of a visualization for a cognitively complex task. These theories, though fundamental, do not address how users think or how visualizations can be applied as an extension to an individual's cognitive ability.

Clearly, we all think differently. There are aspects about you that differentiate you from everyone else. Your experiences, personality and cognitive abilities influence your approach to solving a task and your understanding of a problem domain. In cognitive psychology, researchers have shown that such individual differences can significantly impact a user's dexterity using an interface or a tool to solve problems.

Visualization users differ greatly in experiences, backgrounds, personalities and cognitive abilities, yet visualizations, like much other software, continue to be designed for a single ideal user. It would be clearly impractical to design each visualization for an individual user. However, knowledge of broad differences between user groups could be used to guide design for specific domains and to suggest multiple analysis modes or customization options in a single system. There has recently emerged a new and promising area of research that takes an opposing approach to the traditional method of "one size fits all" design. This research suggests that it is the individual users' cognitive style as much as the visual design that determines the value of a vi-

sualization. Moreover, these individual differences appear to be more pronounced in more complex tasks.

Although still at an early stage, these findings suggest that visualization should not be studied in a vacuum, but in the context of differences among its users. In this viewpoint, we discuss existing work on how cognitive abilities and personality factors affect visualization use, and what still needs to be done in order to advance this field of study. Based on our own experiences in studying individual differences, we argue that current visualization theory lacks the necessary tools to analyze which factors of a design lead to differences in user behavior. Developing this understanding would make it possible to study visualization from the perspective of how an analysis process arises from the interaction between a user and a system. This in turn could lead to a shift in how we evaluate and design visualizations for different user groups, tasks, and domains. In order for this to happen, we must first understand what individual factors affect the use of visualizations.

2 COGNITIVE FACTORS IN VISUALIZATION

Cognitive factors such as spatial ability, verbal ability, and working memory capacity vary substantially between individuals, and can affect reasoning in many different ways. Spatial and perceptual abilities in particular have been shown to affect how well users can perform several different tasks in a visualization system. Velez et al. [17] first showed that a number of these abilities, including spatial orientation, spatial visualizaiton, visual memory, and perceptual speed, affect accuracy and response time on a task involving the comprehension of 3D views similar those found in scientific visualization applications. While this is perhaps unsurprising, subsequent work has shown that these abilities can affect more abstract 2D visualization tasks as well.

Perceptual abilities include basic visual proficiencies such as scanning speed and visual memory capacity. For example, Conati and McLaren [5] found that perceptual speed, which measures the speed at which a person can compare two figures, correlates with a user's accuracy at information retrieval tasks in one of two visualization systems: a star graph and a heatmap-like view. Users with high perceptual speed performed better with the heatmap-like view than the star graph on a comparison task, and vice versa. The authors found that this was only true for one of the tasks they studied, one in which participants were asked to compare differences in change over time between two scenarios at a global level. This was perhaps the most complex question they asked, as most of the others ask the participant to retrieve or compare a specific variable value. The fact that a more complex inferential task is the most susceptible to individual differences is notable, as we found similar effects in our own research.

Spatial ability can be measured by a variety of different tests, which may express different aspects of this factor. In general, however, it refers to the ability to accurately reproduce and manipulate spatial configurations in working memory. In one study of the role of spatial ability in visualization use, Cohen and Hegarty [4] found that a user's spatial ability affects the degree to which interacting with an animated visualization helps him or her perform a mental rotation task. Participants were asked to draw cross-sections of a complex 3D object. They were able to control two animated rotations of the object in order to complete the task. Participants with high spatial ability produced more accurate cross-sections and used the visualizations more, while those

---
[*]e-mail: cziemki@cs.brown.edu
[†]e-mail: alvittao@cs.tufts.edu
[‡]e-mail: rcrous01@cs.tufts.edu
[§]e-mail: krysta.chauncey@tufts.edu
[¶]e-mail:sarasu@google.com
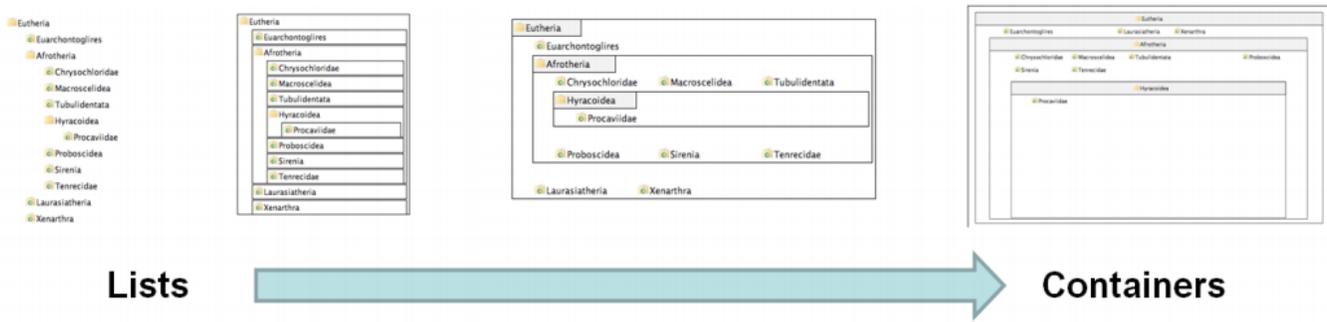[‖]e-mail: remco@cs.tufts.edu

Fig. 1. The four visualizations used in our study on locus of control [19]. These views were designed to vary systematically from a more list-like metaphor to a container-based metaphor. We found that participants with an internal locus of control performed progressively worse (in both accuracy and speed in solving complex analytical tasks) as the visual metaphor shifted from lists to containers. Those with an external locus of control were adept with all views, but especially with the most container-like view.

with low spatial ability rarely discovered the best view from which to create the cross-section.

Similarly, Chen and Czerwinski [3] found a relationship between spatial ability and the visual search strategies users employed in a network visualization. Participants viewed an interactive node-link visualization of a paper citation network and were given tasks to find papers on specific topics. Spatial ability was positively correlated with search task performance in general, and also predicted the use of a better navigation strategy. Low spatial ability participants were more likely to click through every node in a cluster even after determining that the cluster was irrelevant to the target topic, while participants with high spatial ability pursued a more hierarchical strategy in which they jumped from cluster to cluster until they found a likely neighborhood.

Notably, two of these findings show differences in spatial ability affecting not just overall performance, but the way users approach a task. In the cross-section study, high spatial ability participants were more likely to seek out an optimal view for cross-sectioning, and in the network visualization study, they employed a more hierarchical search strategy. The use of different strategies by users with different cognitive profiles suggests that, when user characteristics vary, there is no single way for a visualization to best support a given task. If people with varying cognitive abilities employ different strategies for the same task, a visualization designed for that task must take this into account to be effective.

Further research is needed to elaborate on the implications of these studies. However, they suggest that at least some factors of cognitive ability affect the strategies people use in a visualization task. What remains is to fully characterize these strategies and to be able to predict when differences will arise. While spatial ability is perhaps the natural first individual difference to study in visualization, it is not the only one that has an effect. As Yi [18] has recently argued, a better understanding of individual factors beyond basic spatial ability may be necessary to understand the variability in visualization evaluation. Yi suggests several further factors for study, including visual literacy and personality factors such as field independence and openness to experience. While more work is needed in this area, several experimental results already show that differences in user personality can influence visualization use in significant ways.

3 PERSONALITY FACTORS IN VISUALIZATION

Individual differences go beyond basic cognitive abilities to the more subtle differences that make up a user's personality. Such differences may reflect a user's outlook and common behavior patterns, and recent compelling research on personality differences suggests that they may have a significant effect on performance. Some of these studies indicate that these effects are strongest when tasks are complex, providing a window on how users differ in using a visualization to support higher-level reasoning.

Personality psychology is a well-established area of research, making it a useful lens through which to better understand how different users approach visualization tasks. A common model in personality psychology, the Five-Factor Model, categorizes personality traits on five dimensions: extraversion, neuroticism, openness to experience, conscientiousness and agreeableness. An individual can be categorized under these personality traits, and longitudinal studies [12] have shown that these remain consistent throughout adulthood.

There has been some research in human-computer interaction showing that these personality factors are signficantly correlated with a user's preference for visual interface designs. For example, a study by Saati et al. [14] on interface skinning in a music player compared preferences for five "skins", or visual themes, that varied only on the dominant color. These results showed that preference for the blue theme was positively correlated with introversion, the yellow theme was preferred by more conscientious users, and more imaginative users preferred the black theme. While user preference may affect adoption rates for a visualization system, differences in performance are more valuable in understanding how people use visualization.

That examination of the role of personality on performance remains relatively sparse in HCI research reflects a commonly-held, if implicit, assumption that personality is more superficial than cognitive ability, and therefore unlikely to affect much beyond surface reactions to design. Nonetheless, such studies, including those summarized in an early review by Pocius [11] have suggested that personality can affect interface use at a deeper level. For example, this meta-analysis indicated that introversion is consistently positively correlated with both programming ability and performance with computer-assisted instruction tasks.

In visualization specifically, recent work has shown that personality traits can significantly affect complex task performance. For example, Ziemkiewicz and Kosara [20] performed an experiment on how conflicting metaphors affect tree visualization evaluations. By varying verbal and visual metaphors in the evaluation conditions, we studied the extent to which different users slowed down in response to metaphor conflicts. The results showed that users who scored highly on the openness to experience dimension of the Five Factor model were unaffected by conflicting verbal and visual metaphors. This study also found a similar effect for spatial ability, but the two factors were independent of one another. While this study did not directly compare performance on different types of visual design, it does suggest that participants with high openness and spatial ability may have an easier time switching between different design metaphors, such as those found in a multi-view system.

Green et al. [6, 7] studied the use of visual analytics interfaces by users with varying scores on the Five Factor model and a personality dimension called locus of control [13]. Locus of control measures the extent to which a person sees herself as in control of events (*internal* locus of control), as opposed to being controlled by outside factors
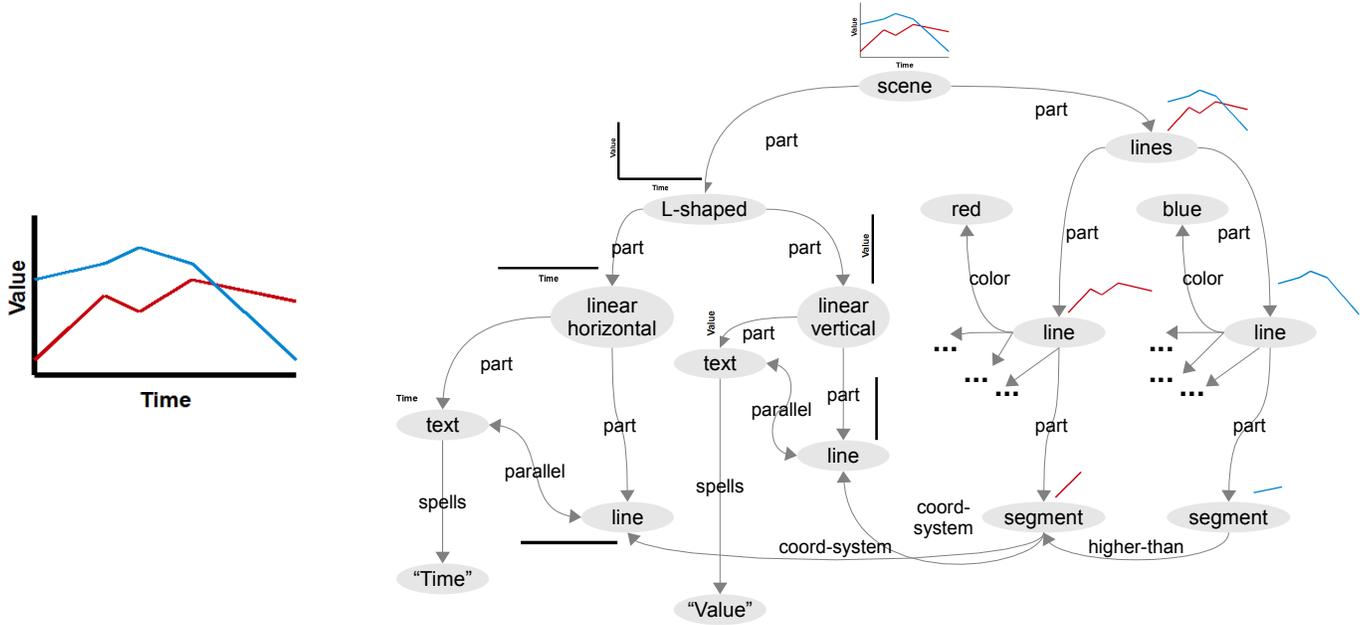
Fig. 2. A simple line chart, and a decomposition of its parts. Adapted from Steven Pinker's "Theory of Graph Comprehension" [10].

(*external* locus of control). The study compared two complex, dissimilar information retrieval systems, a visual analytics system and a web interface with a more list-like view. The authors found that users with a more external locus of control performed better at complex inferential tasks when using the visual analytics interface, and discovered additional correlations between neuroticism and task performance.

Building on Green et al.'s work, we have conducted studies to identify visual elements that appear to be stronger classifiers of users [19]. Our goal was to identify the specific design factors that were responsible for the reported results. Our hypothesis was that the underlying metaphor of the layout was the most significant factor. Therefore, we studied performance on four simple visualizations (Figure 1) that are similar in all aspects except for the overall layout style they use. The purpose of this was to isolate the significant factors in the design of the visualizations at a finer degree of detail than in previous work, which mostly studied real-world visualization tools that differed from one another in many respects.

The four views gradually shift orientation from a list view to one with a containment metaphor. Participants were first measured for locus of control and other personality factors, and then performed a series of search and inferential tasks similar to those used in Green et al. The results showed that, for inferential tasks, participants with an internal or external locus of control performed well on different visualization types, with internal participants showing increased performance as the views became more list-like. In particular, for users with internal locus of control, using a list-like view can produce up to 14.3% increase in accuracy (from 44.4% to 58.7% correct) and 13.6% improvement in response time (from 263 to 227 seconds) when compared to using the containment view. External participants showed less difference in performance overall, but were slightly more adept with the most container-like view than any others. Like Green et al., we found this effect in complex tasks but not simple search tasks.

Results of these studies suggest that personality differences may account for some of the observed individual variability in visualization use. However, this relationship is not a straightforward one. Performance differences based on personality factors appear to manifest for tasks requiring inference and metaphorical reasoning. It is under these cognitively demanding situations that visualization is likely to be the most valuable.

4 RELATING INDIVIDUAL FACTORS TO DESIGN

Prior work has demonstrated effects of cognitive ability and personality differences on visualization use under certain conditions. In order to generalize from these findings, we must isolate the visualization factors and evaluate which ones are helpful or harmful to a user with a given cognitive profile. In turn, we must identify relationships between two primary sets of factors: the cognitive and personality factors that describe the user, and the design and structural factors that describe the visualization. In the case of individual factors, there is a large body of established research in psychology, but little agreement on which are most relevant to visualization. In the case of design factors, there is no real standard language to use when decomposing a visualization. In both cases, we must identify a set of reliable, measurable factors in order to identify useful correlations.

Part of this work is narrowing down which personality factors are most relevant to visualization use. There are already indications of this from previous work: spatial ability appears to be well-established as a factor [17], HCI research points to extraversion [11], and Green et al. [6] have made the discovery that locus of control may be particularly significant for complex visualization use. This work is just the beginning, however. In order to identify which individual factors are relevant to visualization, more studies must be performed both to confirm the factors already found and to investigate new factors. For example, although extraversion was found to be significant in many HCI studies, it has not shown an effect in any of the visualization studies in which it has appeared. Is this due to inherent differences between visualizations and other interfaces, differences in the tasks being studied, or just differences in methodology or study population?

Answering these questions will require both experiments that examine a broader array of individual factors and experiments that study known factors in greater depth. In order to focus this research agenda, a first step may be a formal meta-analysis of the existing findings. This would demonstrate which factors have the most consistent effects, and would be a useful way to find connections between the research on individual differences in visualization and broader HCI research on this subject. Also important is establishing benchmark tasks and datasets so that findings from different studies can be more directly comparable. Progress in this area will not only help to focus the set of factors we study in any given experiment, but also produce information about which aspects of the individual user are significant for visualization. A more difficult question is how to uncover information about which

aspects of a visualization are significant to the individual user.

Much prior work in individual differences in visualization has been case-based and domain-specific. While the researchers in the previously cited studies found individual differences in performance, it is unclear whether those differences are inherent to all visualization use or reliant on aspects of the specifics systems they used. It would be helpful to analyze these systems to discover how they vary from one another. Unfortunately, there does not exist a standardized set of dimensions on which to analyze, let alone synthesize, visual designs.

This proved a challenge when we designed our study on locus of control [19]. Much previous research has aimed to find individual differences in the use of a single visualization system, or in one or more real-world systems that differ from one another in many respects. Our purpose, on the other hand, was to take a known individual difference in performance and attempt to isolate the design factors that influence it. Therefore, we needed to create a set of visualizations that systematically varied on a particular visual quality: in our case, whether the visualization used a containment- or list-driven visual metaphor. We arrived at the solution in Figure 1 through trial and error. Although it was sufficient to isolate the effect we were testing, a more controlled approach and a common language would more quickly advance this field of study. This requires some way to classify a visualization based on its design factors.

There are many high-level classifications of visualization types, but these do not offer the level of detail needed to isolate design variables. Classical visualization taxonomies are often based on data variables, such as dimensionality and data type (e.g., categorical, ordinal, or numerical). For example, Shneiderman's task-by-data-type taxonomy [15] and Tory and Möller's design model taxonomy [16] follow this pattern. Other taxonomies, inspired by Bertin's *Semiology of Graphics* [1], rely on describing the mapping between data variables and visual variables such as color or position. Card and Mackinlay's work [2] on structuring the information visualization design space falls under this category.

While work of this kind is useful in describing the kinds of data a visualization can depict, it is limited when it comes to describing factors of the visual design itself. This makes it difficult to isolate the factors that cause performance differences for varying user types. For example, there is no existing language in visualization theory to easily describe the differences between the visual designs we studied in relation to locus of control. The visualizations are the same in terms of basic visual mapping; the significant design differences are at a structural or metaphorical level. In Card and Mackinlay's descriptions of visual mappings, these properties are generally denoted with an asterisk indicating a special case. There is no systematic way to fit structural design differences into a visual mapping schema.

What is lacking is a usable decomposition of visualization design. In order to correlate individual factors and design factors, we need to know what those design factors are and be able to manipulate them in a controlled way. This means being able to take a single visualization and reliably analyze what components make it up and how they relate to one another. This is an approach closer to that proposed by Steven Pinker [10]. Pinker's goal was to represent charts in a way that would make it possible to model the chart comprehension process computationally. The result is a decomposition into parts that includes both visual mappings, perceptual qualities, and structural elements like axes and labels in a single graph representation (Figure 2).

Pinker based his decomposition theory on simple, static charts, and it is unclear how to extend it for more complex situations like interactive visualizations or multiple linked views. As the existing research shows, it is in these more complex situations that individual differences in visualization user are most likely to arise. Nonetheless, Pinker's work does offer a model for a more comprehensive analysis of the composition of a visualization design. An abstract representation of a visual design produced by decomposition analysis could be measured and analyzed in more quantitative ways, producing metrics that can be usefully correlated to individual personality factors.

Individual factors and visual design are both necessary components in order to explain visualization performance in the context of user differences. However, they are not the only factors in performance. Further research in this area should also consider the effect of data complexity and task type in individual visualization use. For example, we and other researchers have already shown that complex tasks are more likely to be associated with individual differences. Nonetheless, we argue that the first step in this process is understanding significant factors of the user and of the visual representation.

After building a rich taxonomy of design factors that interact with various personality traits and a good understanding of which traits are significant for visualization use, we will be poised to begin running valuable experiments to determine how individual differences affect visualization use. Such experiments could potentially include examinations of how other factors such as data and task complexity play into individual performance differences. This experimental toolkit could also form the basis for deeper questions about how people make sense of visual information under varied circumstances. This is an ambitious long-term research agenda, but its results have the potential to transform our understanding of visualization.

5 CONCLUSION

As we begin to understand the complex relationships among personality, design, and performance, we move toward a fundamental shift in how we approach design of visual interfaces. A goal of such interfaces is to support user thinking. Acknowledging that there is no single, representative user is a critical step toward more effective visualization design. Existing visualization theory has provided extensive knowledge about how to create good visual information designs. However, evaluation results are often ambiguous when two well-designed systems are compared to one another. By understanding the differences between individuals, we may gain the ability to select between good designs to find the best design for a given user. This approach is reinforced by empirical findings, outlined in this article, that individual personality factors affect performance on visualization tasks. Evidence suggests that these performance differences are more apparent in cognitively demanding situations.

Our formal model of visualization must incorporate models of individual users, their personality profiles, and their situational strategies. There are many rich areas for exploration that accompany this ideological shift. For example, because the possible combinations of personality traits are functionally limitless, the process of designing for the user as an individual inherently demands the development and adoption of adaptive interfaces. That is, the interfaces we develop with individual differences in mind should learn about the user as an individual, and adjust themselves to best support the unique combination of personality factors expressed in the user at that time. Such personalized, or adaptive, interfaces are designed to enhance an individual user's strengths and address individual weaknesses, and have been extensively studied in human-computer interaction [9]. Combining this existing work with the knowledge of users we gather in visualization studies will make it possible to tune a visualization interface in accordance with the principles uncovered by this research.

Adapting visualizations to broad classes of users is a valuable strategy for design. However, it is impractical to subject every user of a real-world system to the kind of multiple-choice personality inventories or tests of cognitive ability used in the experiments outlined above. In lieu of laborious tests, we propose to build a model of a user's personality and cognitive ability by analyzing his or her interaction history over time.

While there has been little research in this area to date, exploratory work by Khan et al. [8] shows that significant correlations can be found between interaction measures in a programming task and several personality measures, including those in the Five-Factor model. For example, a negative correlation was found between openness to experience and the number of times a participant switched between windows. As this dimension also showed a positive correlation with the length of time between interaction events, this suggests that more open participants spent more time in each window. Findings like these, extended to visualization-specific tasks, could form the basis for a model of user personality based on tracked interactions. We posit that such a model

of usage patterns will allow us to extrapolate a user's cognitive profile and adapt the visual design accordingly.

While there exists a foundation of individual differences research on which to build, in order to make progress toward adaptive visualizations, more work is needed to put this research in context. We must understand what individual factors are important to visualization use and develop methods for isolating design factors of a visualization. Only with a clear understanding of both can we investigate why users respond differently to different visualizations and apply that information to designing for the user.

The challenges to designing for individuals are great, but the potential benefits make this a challenge worth pursuing. At the individual level, we each stand to benefit from systems that improve our efficiency and accuracy. At the more global scale, many marginalized and traditionally underserved user groups stand to benefit from increased access to visualization systems tailored to them, rather than those designed only for the average user. Finally, this research effort will result in a much deeper understanding of how users make sense of visual information. Visualizations are tools for thinking, and we cannot understand visualization until we understand what people do with those tools. Understanding that there is no one answer to that question is an important step towards truly understanding visualization.

REFERENCES

[1] J. Bertin. *Semiology of Graphics*. University of Wisconsin Press, 1967.
[2] S. K. Card and J. Mackinlay. The structure of the information visualization design space. In *IEEE Symposium on Information Visualization*, pages 92–99, 1997.
[3] C. Chen and M. Czerwinski. Spatial ability and visual navigation: an empirical study. *New Review of Hypermedia and Multimedia*, 3(1):67–89, 1997.
[4] C. A. Cohen and M. Hegarty. Individual differences in use of external visualisations to perform an internal visualisation task. *Applied Cognitive Psychology*, 21:701–711, 2007.
[5] C. Conati and H. Maclaren. Exploring the role of individual differences in information visualization. In *Advanced Visual Interfaces*. ACM Press, 2008.
[6] T. Green and B. Fisher. Towards the personal equation of interaction: The impact of personality factors on visual analytics interface interaction. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on*, pages 203–210. IEEE, 2010.
[7] T. Green, D. Jeong, and B. Fisher. Using personality factors to predict interface learning performance. In *hicss*, pages 1–10. IEEE Computer Society, 1899.
[8] I. A. Khan, W.-P. Brinkman, N. Fine, and R. M. Hierons. Measuring personality from keyboard and mouse use. In *Proceedings of the European Conference on Cognitive Ergonomics*, 2008.
[9] P. Langley. User modeling in adaptive interfaces. In *Proceedings of User Modeling*, pages 357–370, 1999.
[10] 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.
[11] K. E. Pocius. Personality factors in human-computer interaction: A review of the literature. *Computers in Human Behavior*, 7:103–135, 1991.
[12] B. W. Roberts and W. F. DelVecchio. The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, 126(1):3–25, 2000.
[13] J. B. Rotter. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80(609), 1966.
[14] B. Saati, M. Salem, and W.-P. Brinkman. Towards customized user interface skins: investigating user personality and skin colour. In *HCI 2005*, pages 89–93, 2005.
[15] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *IEEE Symposium on Visual Languages*, pages 336–343, 1996.
[16] M. Tory and T. Möller. Rethinking visualization: A high-level taxonomy. In *IEEE Symposium on Information Visualization*, pages 151–158, 2004.
[17] M. C. Velez, D. Silver, and M. Tremaine. Understanding visualization through spatial ability differences. In *IEEE Visualization*, pages 511–518, 2005.
[18] J. S. Yi. Implications of individual differences on evaluating information visualization techniques. In *Proceedings of the BELIV Workshop*, 2010.
[19] C. Ziemkiewicz, R. J. Crouser, A. R. Yauilla, S. L. Su, W. Ribarsky, and R. Chang. How locus of control influences compatibility with visualization style. In *IEEE Visual Analytics Science and Technology*, 2011.
[20] C. Ziemkiewicz and R. Kosara. Preconceptions and individual differences in understanding visual metaphors. *Computer Graphics Forum*, 28(3):911–918, 2009. Proceedings EuroVis.