CGV: HCC: Small: Modeling the Scientific Process for Visualization Evaluation

Caroline Ziemkiewicz (PI) and David Laidlaw (Co-PI), Brown University Department of Computer Science

Project Summary

The objective of this research is to develop a cognitive model of a user exploring scientific data in a multi-windowed data visualization system. The purpose of building such a model is to predict how a users process will be affected by different visual configurations, in order to improve evaluation and design of such systems. This model will be based on the theory that users find information and solve problems in such a system by actively constructing a mental model that combines information from different views.

To accomplish this objective, the researchers will first conduct a series of user studies to observe how scientific researchers employ visualization systems in their work. The focus of these studies will be on recording the mental steps scientists take in pursuing a research task, and how those steps are reflected in their interactions with a visualization. Interaction logging, eye-tracking, video recording, and other measurements will be taken to record this information. Initial data from these observations will be used to guide the design of a cognitive model, using a cognitive architecture such as ACT-R/PM or CPM-GOMS, to predict how the layout of a multi-view visualization system affects this translation between mental process and interaction.

This model will be tested in a subsequent controlled study that compares its predictions about user behavior in varying visual configurations to that of participants. Once a final, validated model has been developed, it will be used to produce guidelines about how to design multi-view visualization systems to optimize for user comprehension and task performance. The researchers will deploy systems redesigned according to these guidelines to domain collaborators to test their efficacy in a practical setting.

Intellectual Merits. Understanding how plans and goals are translated into interaction with a visualization system would be invaluable in the field of visualization as well as in broader human-computer interaction. Guidelines about visualization system design are often based on intuition and anecdotal evidence. Having a concrete model backed by empirical evidence would improve the confidence of these guidelines and can be expected to suggest new ones. Cognitive modeling of a user's process has rarely been employed in visualization research. The unique difficulties of evaluating visualization systems suggests the need for a new methodology for assessing visualizations and understanding why and how some work better than others. The proposed research is an attempt to introduce such a methodology. The investigators combine expertise in scientific visualization and cognitive aspects of visualization, and have access to long-standing collaborations with scientific researchers in varying domains, making them well-positioned to execute this work.

Broader Impacts. The proposed research is highly user-centric and will involve frequent collaboration with domain scientists. In addition to contributing to the theory of visualization and human-computer interaction, this can also be expected to directly improve the design of visualizations being used by these collaborators. This will also have the benefit of exposing students in visualization to the requirements of real users. The results of each stage of this work will be published in appropriate conferences and journals in the visualization and human-computer interaction domains. In addition, the code of the final model will be released to the public under an open-source license.

Keywords: Scientific visualization, user modeling, human-computer interaction, cognitive science

CGV: HCC: Small: Modeling the Scientific Process for Visualization Evaluation

Caroline Ziemkiewicz (PI) and David Laidlaw (Co-PI), Brown University Department of Computer Science

1 Overview and Objectives

Complex data visualization systems are becoming increasingly ubiquitous in many areas of study, yet evaluations of these systems are often unreliable and inconsistent. Part of the problem lies in the mismatch between traditional evaluation methods and the types of tasks supported by visualization tools. Using a data visualization tool is typically an open-ended reasoning process that makes use of an external visual representation as a cognitive aid. Without a concrete understanding of how a user makes sense of this visual information in a reasoning process, it will remain difficult to assess how well a visualization is performing its job. There have been several successful computational models of visualization use at the perceptual level, which have led to improved guidelines for design factors such as size scales and color maps. However, existing models of the cognitive level of visualization tend to be loose frameworks, and are rarely specific enough to make clear predictions about the effectiveness of a specific visualization design. Computational models that can make such predictions have the potential to transform visualization evaluation and design and create real progress towards a rigorous science of visualization.

Proposed approach. The long-term goal of this research is to advance this science by contributing to a set of cognitive models that predict and explain user behavior with a visualization tool. To move forward with this goal, the objective of the proposed work is to develop and validate a cognitive model of a scientific user analyzing their data in the context of a multi-view visualization system. Using an interdisciplinary approach that combines visualization and graphics research with cognitive science, we plan to create this model by first *observing* the scientific user, then *building and validating* the model, and finally *applying* the model to evaluation and design.

Our working hypothesis is that scientific users typically approach a data analysis task with the central goal of forming a coherent mental model of their data that can answer certain types of research questions or drive hypothesis formation about a specific topic. Therefore, the primary process of using a multi-view visualization system is an iterative one in which multiple data facets and data representations, represented by visualization views, are incorporated into a user's mental model. We hypothesize that the type of information present in these views, how they are represented and arranged, and the visual and interactive links between them all contribute to the accuracy and efficiency with which this mental model is built. This hypothesis is based in Dr. Ziemkiewicz's prior work showing that visual structure can significantly affect data comprehension in visualization use [29, 30]. When validated, this model could be used to automatically evaluate the effectiveness of a multi-view visualization design and provide guidelines for the design of future systems.

Investigator expertise. The investigators are highly suited to studying and executing this vital problem. Dr. Ziemkiewicz provides a unique background in the study of how visual structure affects visualization use and understanding, while Dr. Laidlaw is a leader in producing successful scientific visualization research and applications and provides established collaborations with research groups in a variety of scientific domains, including brain research, proteomics, and biomechanics.

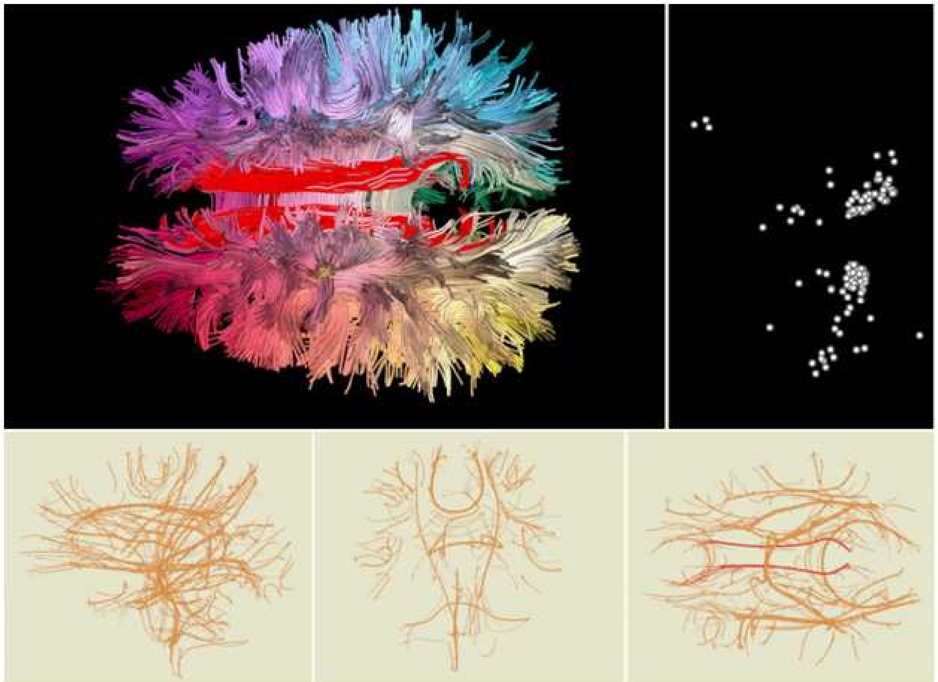

Figure 1: A visualization of diffusion tensor imaging (DTI) data [12], developed in collaboration with brain scientists at Brown University, and one of the scientific visualization systems that we propose to study. Complex mult-view visualization systems of this kind are both critical for scientific data analysis and difficult to evaluate and design.

Research plan. The first goal of the proposed research is to thoroughly study the reasoning process of scientists using visualization tools in practice. Through observation and task analysis, we will identify the typical steps taken by scientists during data analysis tasks. We will then use interaction logging, eye-tracking, and other measurements to study how those steps are reflected (or not reflected) in their use of domain-specific data visualization systems. The intention of this portion is to test our initially hypothesized model of the scientific visualization reasoning process and modify that model as needed to reflect real user behavior. The second goal will be to implement this process as a concrete computational model that takes into account the design parameters of the visualization to predict the steps needed to accomplish a given analysis task using that system. We will then test these predictions by comparing them to actual user behavior in a controlled experiment. These two steps will be iterated as needed to produce a model that can produce reliable predictions of user activity. This final model will be then used to evaluate the efficiency of existing visualization systems in use by domain experts and to suggest guidelines for improving these designs. The final goal of this research will then be to deploy these improved designs to domain scientists and evaluate their effectiveness in practice.

The proposed research represents a significant break from traditional visualization evaluation methods. Rather than attempting to extract design recommendations from response times and error rates measured in side-by-side comparisons of visualization systems, our proposed method begins with a theory about why two visualization systems might be different and rigorously tests the predictions of that theory both in the lab and in practice. By carefully grounding theory in observation, implementation, and validation, this approach has the potential to tie visualization theory more closely to visualization practice. We expect this research to produce a generally useful body of knowledge about the reasoning and interaction process of scientific visualization users, a computational model that can be used for evaluation, and a set of validated design guidelines for building effective multi-view visualization systems. Taken together, we expect these outcomes to lead to a greater understanding of how people use visualization as well as tangible improvements in visualization design. Doing so will advance the science of visualization and promote a tighter coupling between theory and practice.

2 Expected Significance of Model Development

Visualization is growing in widespread use and is increasingly being applied to critical domains such as scientific research, intelligence analysis, and financial analysis. Applications in these domains frequently employ multiple views to convey complex information and multiple data facets, producing visually dense interfaces that can present data that is both massive in size and heterogeneous in nature. At the same time, evaluation methods that can reliably distinguish the good from the bad among these applications remain underdeveloped [26]. Since visualization tools are often used for ill-defined, open-ended analysis tasks, traditional usability testing methods that focus on the speed and accuracy of task completion are less applicable. This may partly explain why results are so frequently inconsistent and difficult to interpret across evaluation studies of visualization tools [4]. These inconsistencies suggest that current evaluation methodologies employ a faulty or incomplete model of what makes a successful visualization. Therefore, there is a need for methods that can not only distinguish successful visualization designs from unsuccessful ones but can also explain *why* they are successful.

The proposed research will contribute such a method in the form of a cognitive model that predicts and explains the analysis steps taken by a user given a task and the design parameters of a multi-view visualization system. Instead of using time and errors as a coarse measurement of a user's ability to reason with a visualization, this method would produce a more fine-grained assessment of what kind of user behavior a visualization design is likely to encourage. A successful cognitive model of this kind would represent a novel direction for visualization evaluation as well as a significant advance in the science of visualization. Benefits would include the ability to base design guidelines for multi-view visualization design on theoretically grounded empirical results, increasing confidence in future designs. It would also offer a method for other researchers and designers to assess their own designs by downloading and applying our model.

Iimproving this common type of visualization design would have an immediate impact on the many fields that use such applications, including the sciences and visual analytics applications. More broadly, this would offer a demonstration that cognitive modeling can be used for more nuanced visualization evaluation, potentially inspiring further research in this area and a collection of models specialized for various problems. Finally, it would advance the science of visualization by offering a concrete, validated theory for how users make sense of information in multi-view visualization systems.

3 The Challenge of Multiple Coordinate Views

In this work, we focus on scientific visualizations with mulitple coordinated views (coordinated multi-views), such as the one shown in Figure 1. Visualization systems of this kind are very common in complex data analysis tasks, both in scientific domains and elsewhere. The goal of such systems is to simultaneously present different aspects of a dataset, comparisons of different datasets, or connections between separate but related data. Frequently, these systems use a coordinated multi-view interaction style, in which changes in one view (such as highlighting, selection, and filtering) are reflected in the others. North and Shneiderman [23] provide a thorough description of this interaction style as part of their snap-together visualization toolkit.

This type of visualization design is intended to help a user gain a coherent understanding of how different facets of a big picture view of information relate to one another while also presenting specific parts of the overall picture individually. While this is a powerful analysis technique, it also presents many unique challenges for design and evaluation. The science of visualization is still in its early stages, and evaluating a single visual representation can be a difficult task in and of itself. When a viewer is asked to synthesize several separate representations into a coherent picture of data, this design problem is significantly compounded. For some tasks, the user will need to rapidly switch attention between different views, requiring constant orientation to new information. In addition, instead of learning a single visual encoding, the viewer often must learn several encodings and all the correspondences among them.

Furthermore, we argue that this is increase in complexity also carries significant cognitive costs. We adopt the view that user of a visualization constructs a mental model of data based on their interpretation of the visualization's information structure. The purpose of this mental model is to record relationships between parts of the data in some kind of cohesive structure. If the visualization implies a tree structure, this mental model will be constructed around hierarchical relationships between data items; if the visualization resembles a map, the mental model will record spatial relationships like distance and clustering. With each new pair of visual representations on a screen, a user needs to find a way to translate between the mental models of data implied in each. When there are four or five views on a screen, this could force the user to generate a complex mental model that somehow combines aspects of each view, increasing the possibility of errors and confusion.

Although these issues make multi-view visualization design difficult, the promise of coordinated multi-views is such that solving these problems should be a priority for visualization research. Wang Baldonado et al. [28] have proposed a set of design guidelines for multi-view visualizations based on their own design experiences, but there remains a need for more empirically grounded guidelines that explain not only what should be done in designing these systems, but why.

4 Cognitive Modeling

To approach the problem of understanding how people make sense of multi-view visualizations, we propose to use cognitive modeling. Cognitive modeling is the process of implementing cognitive science theories as computational models. At its simplest, this can mean developing a mathematical equation that approximates some process. For example, Plumlee and Ware's model of visual comparison [25], discussed in Section 5 is of this nature. There are also specialized toolsets known as cognitive architectures that provide a coding language and interpretation environment for modeling. These architectures automatically simulate certain general mental processes so that the modeler can focus on the specific process she is trying to explain. How these processes are simulated varies between different cognitive architectures, making each specialized for models with particular goals and assumptions. An example of a cognitive architecture that may be used in this work is ACT-R (Adaptive Control of Thought–Rational) [1], which simulates memory storage and retrieval as well as some simple goal representation. Other cognitive architectures include Soar [11], which represents task decomposition in pursuit of a goal; EPIC (Executive-Process/Interactive Control) [15], which is specialized for user interaction modeling with a computer interface; and CPM-GOMS (Cognitive Perceptual Model - Goals, Operators, Methods, and Selection rules) [14], which focuses on representing parallel perceptual and cognitive processes for task efficiency analysis.

Models built using these architectures can be used to express a cognitive theory in a concrete way. Typically, these models will take some kind of input in the form of an abstract description of a problem, and then generate the steps a person would take to solve that problem. The nature of these steps depends on the nature of the architecture. In reasoning-driven architectures like ACT-R and Soar, the steps will typically represent a sequence of mental states, defined in terms of the current goal or subgoal and the contents of working memory. Architectures like EPIC that are specialized for user interface analysis might produce a sequence of user actions, such as keystrokes or menu selections. By associating an average time span with these actions, the modeler can then produce a prediction of the total time to execute a particular task. A GOMS analysis of this kind was used in Project Ernestine [9] to create a more efficient software for telephone operators, saving the NYNEX phone company $2 million dollars a year.

Examples like this demonstrate that cognitive modeling is a powerful tool not only for making theoretical concepts concrete and testable, but also for producing actionable evaluation metrics. Since the long-term goal of this research is to build theories of visualization use that can be applied to practice as well as to advancing science, building a cognitive model is an ideal approach. It is nonetheless vitally importnat that this model be grounded in observation of user behavior and thoroughly validated. The history of modeling in visualization evaluation and design suggests that the most successful cognitive models in this domain are those that are most clearly based in empirical data.

5 Background and Relation to Long-Term Research Goals

This work addresses an unmet need in visualization research for grounded and usable theories about how users make sense of visual information in a multi-view system. Visualization theory so far has produced several working models of how visual marks are interpreted at a perceptual level, which offers a similar approach to the work being proposed. Some of these models make use of Bertin's *Semiology of Graphics* [3], which first defined "marks" as the most basic elements of graphical encoding and defined possible mappings between types of data dimensions (nominal, ordinal, or numerical) and "retinal variables" of graphical marks (size, color, position, etc.). One of the most notable extensions of this work is the research performed by Cleveland and McGill [5] that took Bertins basic taxonomy of visual variables and measured how finely viewers could discriminate between them in a graph-reading context. This produced a ranking of visual variables in terms of their discriminability, and concrete guidelines for which variables could be used for numerical data (e.g., length and position) and which should be reserved for nominal data (e.g., color). This provides a grounded, usable metric for one element of visualization design. Indeed, Mackinlay's APT (A Presentation Tool) [20] proposed using these mappings to automatically generate the best visual encoding for a given set of data, a system which has been implemented in the commercial visual analytics software Tableau [21].

A similar philosophy has been applied to research on choosing pereceptually consistent color maps [2]. By focusing on specific visual variables, researchers have also developed mathematical models of texture discrimination [10] and symbol size discrimination [16]. Models such as these can be used to set limits on the scales used by visualization designers when employing these variables in a data mapping. By modeling perceptual processes and producing design guidelines based on their predictions, this body of research shows the power of a working visualization model for improving design and evaluation. The proposed work would build on such research by extending it to a cognitive domain rather than a strictly perceptual one.

Models that take a more cognitive perspective on the interpretation of visual information are rarer, but do exist and form an important foundation for the proposed work. Pinker's theory of graph comprehension [24] describes in detail the operations taken by a viewer of an information graphic, based on the type of visual mappings used in the graph and the particular query that a user is attempting to answer with it. Using a cognitive modeling language which is a variation on ACT-R, Lohse [19] constructed a computational model based on Pinker's theory as a test of its validity. This is an excellent foundation for a general-purpose graph comprehension model, but it has several limitations that prevent it from being immediately applicable. For one, it refers only to static graphs, while modern visualization systems use interaction as a core component. For another, Lohse's results are not strong enough to convincingly argue that this model sufficiently captures the process at work. The proposed work aims to increase the predictive power of a similar model both by focusing on a more specific subset of visualization problems and by grounding model design in close observation of users.

A recent model of a more specific visualization process is the work on comparing zooming versus focus-and-context views by Plumlee and Ware [25]. The authors designed a mathematical model to predict how quickly and accurately a user will be able to make comparisons between detailed views in a geospatial visualization, based on whether those views are reached by a zooming operation or by generating detail windows and keeping the overall map visible. Their model proved to be successful at making predictions and explaining why focus-and-contex views are more efficient for comparison. This work is closely related to the proposed work in that it involves modeling comparisons between multiple windows of a visualization system, but deals with visual comparisons rather than a broader analysis task. These models demonstrate that it is possible to make sense of higher-level visualization understanding and use through computational modeling, which provides evidence for the validity of our proposed approach.

The proposed work takes the working hypothesis that exploratory data analysis with a visualization system is largely a process of adding disparate information to a tentative mental model of data. This draws on recent research that considers a visualization as a component in a reasoning process. One prominent example is a model by van Wijk [27] based on cost-benefit analysis. In this model, the purpose of visualization is to increase a user's knowledge (ΔK) by expressing data in an image. This model has the benefit of including interaction, as opposed to the more static model proposed by Pinker, but it models the process at a very coarse level. It is not clear how this could be used to explicitly guide design or evaluation, a major goal of the proposed work.

Other high-level models of this kind take an approach inspired by cognitive science. Liu et al. [17] have proposed a theory of visualization based on distributed cognition. In this formulation, a visualization system is an external representation that forms part of a user's knowledge base and thinking process when solving a problem. This approach views the user and the visualization as parts of an overall system that together make sense of data. Following up on this work, Liu and Stasko [18] argued for mental models as the basis for a theory of visualization. adding more detail to what happens in the user's head during visualization use. In this view, visualization is largely a process of adding to an internal mental model of information by incorporating parts of the visualization. This perspective draws heavily on work from cognitive science that theorizes mental models as central to reasoning. Indeed, there is evidence from cognitive science [7] that mental models are a valid way to describe how people make sense of diagrams and visualizations. Dr. Ziemkiewicz's prior work [29] adds to this evidence by demonstrating that a correspondence between visual metaphors and a user's internal metaphors for data produces more efficient visualization use. This general body of work provides some provocative ideas about how people make sense of visual information, but as of yet has not produced a model that can be obviously proven or falsified.

The proposed work represents a more concrete step forward for this area of research. The long-term goal of such research is to contribute to a grounded science of visualization by building theories that can genuinely explain and predict user interaction with a visualization system. Valid theories are vital for the continuing development of visualization as a scientific discipline and for the ability to guide visualization research in useful directions. Dr. Ziemkiewicz has previously contributed to this line of research by extensively studying how visual structure is interpreted by users. The proposed research extends this work in a more applied direction by investigating how multi-view visualization layouts guide a user's development of a mental model of their data. When completed, the proposed model is expected to serve as a first addition to a collection of validated cognitive models for visualization evaluation. Building this collection over time will provide vital tools for future evaluation studies and act as a library of concrete visualization theories that can be studied, compared, and combined for future theoretical research.

6 Plan of Research

To contribute to these long-term goals, we have designed the following plan of research as a tight coupling of theory with execution and validation. Our general approach is to first observe users in order to ground this theory, then to implement that theory in the form of a computational model and validate it by comparing it against empirical data in a controlled study. Once we have built a validated model, we plan to apply it to evaluation and design to demonstrate its practical utility. The following timetable and plan of research details the steps we will take to execute this approach.

6.1 Phase 1: Observing the Scientific Visualization User (six months)

A solid understanding of the analysis process of scientific users is needed to inform a general model of the kind we intend to build. Therefore, the objective of this portion of the research is to have a collection of data and knowledge on which to build our initial model. Our working hypothesis is that using a multi-view visualization system is an iterative process in which multiple data facets and data representations, represented by visualization windows, are incorporated into a user's mental model. The type of information present in these views, how they are represented and arranged, and the visual and interactive links between them all contribute to the accuracy and efficiency with which this mental model is built. We will test this hypothesis by observing the stated reasoning process of working scientists and how they use a visualization system to execute that process. In doing so, we will leverage the existing working relationships between the Brown Visualization Research Lab and several domain-specific scientific research labs at Brown University. These projects include cognitive neuroscience work with The Badre Lab; diffusion tensor imaging visualization with Dr. Steven Correia [8]; cell signaling proteomics with the Salomon Research Group [13]; and wrist biomechanics with researchers at the Alpert Medical School [22]. The Visualization Research Lab has developed or partially developed visualization systems with each of these groups, so there are already active systems in the field. By beginning our research in this way, we will ensure that the subsequent models are grounded in reality and reflective of actual user behavior. We expect this portion of our research to produce a rich body of knowledge that reflects general thought processes and interaction styles of varying scientific visualization users, explaining what they do with a visualization system and why. The following subsection details how we intend to build this body of knowledge.

6.1.1 Research Design

The objective of this phase is to closely observe how users employ these visualizations in pursuit of a data analysis task. The first aim of this work, therefore, will be to define these analysis tasks, independently of specific visualization systems, and how the tasks fit into an overall scientific process. The purpose of this portion is to produce an "ideal" analysis process without overly biasing the analysis in favor of our existing visualization designs.

- **Interviews with domain scientists.** We will conduct detailed interviews with as many members of the scientific research labs as possible (a minimum of ten). These interviews are expected to discover what the scientists intend to do with their data analysis, regardless of whether the visualization allows them to do that or not. We will ask the scientists what their analysis questions are, what problem-solving process they go through when trying to answer those questions, what obstacles they tend to encounter, and so on. This portion of the research is currently in progress.
- **Observation of analysis work.** Where possible, we will schedule sessions where one of the members of our research team visits a domain scientist and observes actual data analysis work that does not employ a visualization.
- **Task analysis.** Based on the findings from the interviews and observations, we will perform a task analysis of the analytical work done by these domain scientists. We will first classify the tasks we observed into types if necessary, with the goal of having the fewest possible coherent task types. We will then use task analysis methods to decompose the tasks into steps and identify the subgoals of each step.

The second goal of this phase is to observe and measure visualization use and relate it to elements of the task analysis.

- **Experimental setup.** In collaboration with each of the domain scientists, we will identify a task of the type for which they are likely to use a visualization, using novel or invented data so that the scientist will not have performed this particular analysis before. We will then set up a session for the scientist to perform the task while being tracked.

- **Visualization use logging.** We will employ an eye-tracking device, interaction logging software, screen capture software, and video recording to track the user during their performance of the analysis task. Immediately afterwards, we will present the video to the participant for a post-hoc think-aloud protocol in which they describe their thought process while viewing the video of themselves. This is intended to provide a reasonably accurate recording of the participants' thoughts during the study without disrupting their analysis process at the time.

- **Analyzing the results.** Analysis of the study data will focus on the sequence of eye movement, particularly when participants switch from one view of the multi-view system to another, as well as the sequence of interactions. This information will be combined into a *visualization use sequence*, a timetable and list of events that represent changes in attention and interaction mode. We will also record any errors or obstacles that arise in the course of the analysis process.

- **Comparing results to task analysis.** We will compare the visual use sequences to both the general task analysis we produced earlier and to the participant's own post-task description of their thought process. The goal will be to find the points in which the visualization use sequence follows logically from the steps of the analysis process, and points where the two do not seem to align. This comparison will form the basis of the initial model design in Phase 2.

6.1.2 Expected Outcomes

This phase of the research is expected to produce the following outcomes:

- A task analysis of the analytical activity undertaken by scientific users in a variety of domains.

- A body of data describing scientific visualization analysis with coordinated multi-views in practice, including quantitative measures of interaction sequences and eyegaze behavior as well as qualitative information gathered from interviews, video recording, and post-experiment talk-aloud protocols.

- A correspondence between mental steps in a data analysis task and interaction and interaction and attention-switching in a multi-view visualization system.

We expect this phase of the research to take about six months, including time to design our observational methods, interview our collaborators, set up tracking equipment and interaction logging software, and schedule visits with the outside research labs.

6.2 Phase 2: Building and Validating a Cognitive Model (15 months)

The primary goal of this phase of the research is to take the body of data generated in Phase 1 and explain it using a coherent computational model. Our working hypothesis is that this model will describe how a user builds a mental model by combining information from different views. However, if this hypothesis does not appear to be a good description of our measurements and observations from Phase 1, we will formulate a revised hypothesis based on our empirical findings. Our planned approach is to use a cognitive modeling architecture (Section 4) which can combine models of perceptual behavior with models of higher-level cognitive functions. The purpose of this model is to predict a probable sequence of interaction and viewing steps given a visualization configuration and a type of analysis task. This portion of the research will be iterated as necessary to produce a model that makes valid predictions about user behavior.

6.2.1 Research Design

The first step in this iterative process is to build an initial computational model, based on our working theory about how a scientist's reasoning process is reflected in interaction with a multi-view visualization.

- **Generating a working hypothesis.** By modifying our initial theory based on the findings from Phase 1, we will generate a basic hypothesis about what the analysis process is doing and why. This hypothesis will act as a general guide to cognitive model development.
- **Identifying visualization parameters.** Using this hypothesis and our observations from Phase 1 as a guide, we will identify design parameters of multi-view visualization systems that are likely to be significant inputs to our model. These parameters may include the number of views, the spatial arrangement of views on the screen, whether interaction within the views is coordinated, how similar they are to other views in appearance or data content, and so on.
- **Sketching a conceptual model.** We will begin with a conceptual design of the cognitive model on paper, which can go through several rounds of revision by the research group.
- **Building a computational model.** We will then develop a cognitive model based on the sketched design, using one of the architectures described in Section 4. This model will take as inputs the design parameters of a visualization and a task decomposition from Phase 1, and generate a prediction of the visualization use sequence needed to execute the task.
- **Comparing predictions to Phase 1 findings.** Once this initial model is built, we will test whether its predictions reasonably approximate the visualization use sequences recorded in Phase 1 when the model is given similar input. If the model produces sequences that clearly contradict the empirical data, we will alter the design accordingly.

When the model seems plausible when compared to our initial data-gathering, we will continue to the next step of testing its predictions in a controlled setting.

- **Experimental setup.** We will begin by designing controlled test visualizations, broadly resembling those used by our collaborators, that only differ from one another on the design parameters that act as input to the computational model. This is intended to clearly isolate these factors in the experiment. Tasks will then be designed that can be performed with these test visualizations and which represent each of the task types identified in Phase 1.
- **Generating model predictions.** These design configurations and task types will be used as input to the computational model in order to generate visualization use sequence predictions for each of the conditions that users will see in the subsequent experiment.
- **User study.** We will then perform a large-scale study of visualization use sequences under the different test conditions. In order to reach a sufficiently large number of participants, this study may include layperson participants, but will also include some of the experts from

the domain research groups to test for any differences between experts and novices. The users will be trained on the visualizations and then asked to perform the tasks across each of the conditions. Their interactions and eye-tracking data will be recorded to determine the sequence of their visualization use. Additionally, the task completion time, error rate, and reported usability of the systems will be recorded, to test how the sequencing relates to traditional measures of user interface performance.

- **Analyzing the results.** We will analyze the data from this study with a focus on whether the recorded use sequences of participants usually match the predictions of the model. We will also examine whether the model's predicted task sequences can help to explain error rates, response times, and reported usability across the study conditions. If the initial model does not appear to make valid predictions about use sequences, we will look into this data for an explanation of why and to determine what kind of information the model does not seem to capture.

6.2.2 Expected Outcomes

If the initial model does not prove to be a valid predictor of user behavior at the end of this phase, an abbreviated version of this phase will be repeated as necessary to produce a valid model. When all iterations are complete, this phase of the research is expected to produce the following outcomes:

- A validated model which can be used to predict a user's interaction and view sequence with a multi-view visualization system.
- A cognitive theory to explain why certain multi-view configurations are better than others.

We expect this step to take two iterations. Initial model-building is expected to take about six months, and the validation study should take three months each to execute. Once there is an existing computational model, subsequent rewrites and studies of the model will be less time-consuming, so the second iteration is expected to take a maximum of six months. Therefore, this step is expected to take about 15 months altogether. Should more than two iterations prove necessary, we will begin performing smaller-scale validation studies to decrease turnaround time.

6.3 Phase 3: Applying the Model to Practice (three months)

The production of a computational model that can be used to predict visualization user behavior is a significant goal in and of itself, as it would embody a theory of visualization that is more concretely grounded than most. However, in order to have the greatest possible impact on visualization research and practice, it must also be proved that this model can be usefully applied to design and evaluation. Our approach to demonstrating this practical utility is to use the model to evaluate and improve the scientific visualizations we analyzed in Phase 1. The purpose of this step is to demonstrate that the model from Phase 2 not only expresses a theory but has practical value in visualization system design.

6.3.1 Research Design

The first step in this phase is to evaluate existing systems using our model to see whether it can usefully differentiate between different designs.

- **Explaining interface issues.** In this step, we will use the model from Phase 2 to produce sequencing predictions for appropriate task types using the parameters of the scientific visualization systems studied in Phase 1. We will then look back at any points of frustration or errors found during the observations of Phase 1 and see whether the model can explain why they are happening. For example, the model may predict that a particular arrangement of views forces the user to frequently switch attention across a large screen distance during task execution.

- **Evaluating outside systems.** As a proof of concept, we will also apply the model to visualizations by other researchers and in other domains that have been used in evaluation studies, to see whether the model can help to explain these evaluation results as well. We will also look for systems which have interaction logging data associated with them for additional validation of predicted visualization use sequences, such as the WireVis financial analytics system [6] previously worked on by Dr. Ziemkiewicz.

If the model can be used to explain issues with existing systems, it should also be able to suggest design improvements that can avoid those issues. The next goal of this phase is to develop design guidelines based on the model.

- **Identifying factors of successful designs.** We will start by comparing the model's predictions about more and less successful visualization configurations. The focus will be on which design parameters predict more efficient use sequences. We will also look for common attributes of the use sequences in the more successful visualizations (that is, those with higher usability ratings and lower error rates overall) and find common areas where the less successful designs are causing users to get stuck or behave suboptimally. Looking to the results from the validation studies in Phase 2 can help to isolate the design factors that produced more or less errors, and why.

- **Creating design guidelines.** Based on the common factors identified in this analysis, we will develop and publish a set of design guidelines for multi-view visualization systems. There are existing guidelines of this kind [28], but they are largely based on intuitive interpretations of usability studies. We expect these guidelines to have the advantage of a solid theoretical grounding and empirical validation.

Finally, we will perform an informal practical validation of these design guidelines by redesigning visualizations from Phase 1 and deploying the new designs to domain scientists.

- **Redesigning visualizations.** We will use the design guidelines to produce new versions of each of the visualization systems studied in Phase 1. We will also use the computational model to confirm that the new designs are predicted to encourage more efficient analysis paths through the data than the old designs.

- **Deploying and testing.** The new versions of the visualizations will be deployed in the research labs of our scientific collaborators. In each case, we will perform a short usability evaluation to compare the old and new versions in terms of user preference, task performance, and errors to confirm that this novel evaluation method does not contradict traditional methods.

6.3.2 Expected Outcomes

We expect this phase of the research to produce the following outcomes:

- A set of design guidelines for multi-view visualization systems that is empirically validated and theoretically grounded.
- A demonstration that computational modeling can be used to improve and evaluate visualization designs in practice.
- Better scientific visualization tools for our collaborators.

This phase is expected to last about three months, including the time to perform evaluations and meta-analysis of other evaluation studies, codify the design guidelines, redesign the scientific visualizations from Phase 1, and deploy these visualizations to our domain collaborators. Altogether, the three phases of our research are expected to last for two years.

7 Results from Prior Support

Over the last five years David Laidlaw's research group has been funded by several NIH awards or subcontracts to NIH awards, continuation funding from several NSF awards and a Keck Foundation grant, one recent NSF instrumentation grant, and a collaborative NSF award (IIS-10-16623). Most of the NIH awards and subcontracts related to developing and applying computational tools for the analysis of diffusion MRI imaging data. Laidlaw's group has published extensively in this area with collaborators from around the world. Another NIH award was for developing computational bioengineering tools to study the human carpus in collaboration with PI Joseph Crisco in the Department of orthopedics. A final NIH subcontract involves creating interactive tools for studying gene expression in the immune system. This is in collaboration with Christophe Benoist at Harvard Medical School. Laidlaw's most recent NSF funding was an MRI award to develop a new virtual reality display instrument at Brown. That project is in year two of four. Finally, he has been a co-PI on Keck and NSF awards to study animal kinematics and dynamics.

The most similar award to the one proposed is titled "GV: Small: Collaborative Research: Supporting Knowledge Discovery through a Scientific Visualization Language." This project, funded for $261,596, began six weeks ago on November 1, 2010, with an estimated end date of October 31, 2013. The purpose of this award is to study how best to design a scientific visualization language for creating visualizations of diffusion MRI data. As this work was very recently awarded, there have not been any publications associated with it yet. At a broad scale, however, in the last five years Dr. Laidlaw's research has led to about 25 refereed journal articles, approximately 10 refereed conference papers, two patent applications, and about 40 conference abstracts or posters. He has been recognized via a number of best poster and best panel awards at conferences and received in 2008 the prestigious IEEE VGTC Visualization Technical Achievement Award.

8 Broader Impacts

8.1 Education and Infrastructure

The budget includes funding for a graduate student research assistant, who will be an integral part of the execution of this plan. This student will learn how to conduct user experiments, formulate theories, and build cognitive models to express theories. They will also have the opportunity to conduct interdisciplinary research in collaboration with scientific experts accross domains. At the same time, this research will provide an opportunity to educate students working in brain science, proteomics, and medical research labs on cutting-edge visualization techniques. This will also serve as an opportunity for these students to examine their own analysis process and find ways to critique the analytical tools they use.

Another educational opportunity is provided by Dr. Laidlaw's graduate-level course on Interdisciplinary Scientific Visualization, in which students are asked to propose and execute a class project in collaboration with scientists at Brown University. As this research offers an ideal opportunity for students to learn how to thoughtfully collaborate with domain researchers in building and evaluating a visualization system, we will suggest portions of this research as one of the projects students can choose from.

This work will improve research infrastructure by deepening collaborative relationships between the Visualization Research Lab and other research groups at Brown and elsewhere. In addition, it will improve the analysis tools used by scientific researchers in these groups, allowing their work to proceed more efficiently and insightfully in the future. Another benefit of this close work with users is in its potential to involve underrepresented groups, particularly women, in visualization research. We strongly believe that one of the best ways to broaden participation in computer science is to demonstrate its creative potential and relevance to real-world problems. By closely collaborating with researchers outside of computer science, especially those in medical and biological fields, the proposed work has the potential of exposing female students in other disciplines to creative and practical applications of advanced computer science research.

8.2 Dissemination and Societal Benefits

Along with deploying our improved visualization designs to our scientific collaborators, we will make a concerted effort to disseminate these results to the visualization research community at large. This will include publishing the results in appropriate research conferences and journals at regular intervals. Some of these results, such as the design guidelines from Phase 3, are expected to have an immediate impact on improving visualization development. Others, such as the theory behind the model developed in Phase 2, will contribute to a more long-term discussion about visualization theory and research directions. We will also make the code for the model itself publicly available, both so it can be used by other researchers and designers and so it can be analyzed by other theorists. In general, the proposed work has a unique potential among visualization theory research to directly impact practice as well as the broader understanding of how visualization works.

Finally, publication of design guidelines that are theoretically and empirically grounded will improve visualization designs across numerous fields. As visualization has expanded in widespread use as well as the importance of its applications, the potential impact of foundational visualization research has expanded as well. Work that contributes to the basic scientific understanding of how a user makes sense of complex visual information can help to make tools more effective in a wide variety of fields, both scientific and otherwise.

References Cited

[1] John R. Anderson, Michael Matessa, and Christian Lebiere. ACT-R: a theory of higher level cognition and its relation to visual attention. *Human-Computer Interaction*, 12(4):439–462, 1997.

[2] L. D. Bergman, B. E. Rogowitz, and L. A. Treinish. A rule-based tool for assisting colormap selection. In *Proceedings IEEE Visualization*, 1995.

[3] Jacques Bertin. *Semiology of Graphics*. Univ. of Wisconsin Press, 1967.

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

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

[6] Wenwen Dou, Dong Hyun Jeong, Felusia Stukes, William Ribarsky, Heather Richter Lipford, and Remco Chang. Recovering reasoning processes from user interactions. *IEEE Computer Graphics and Applications*, 29(3):52–61, 2009.

[7] Dedre Gentner and Donald R. Gentner. *Mental Models*, chapter Flowing Waters or Teeming Crowds: Mental Models of Electricity, pages 99–129. Lawrence Erlbaum Associates, 1983.

[8] Assawin Gongvatana, Ronald A. Cohen, Stephen Correia, Kathryn N. Devlin, Jadrian Miles, Uraina S. Clark, Michelle Westbrook, George Hana, Hakmook Kang, Hernando Ombao, Bradford Davia, David H. Laidlaw, and Karen Tashima. Impact of hepatitis C and HIV coinfection on cerebral white matter integrity. *Neurology*, 2010. In Review `http://vis.cs.brown.edu/docs/pdf/bib/g/Gongvatana-2010-IHC.pdf.html`.

[9] Wayne D. Gray and Bonnie E. John. Project Ernestine: validating a GOMS analysis for predicting and explaining real-world task performance. *Human-Computer Interaction*, 8(3):237–309, 1993.

[10] Danny Holten, J. J. van Wijk, and J.-B. Martens. A perceptually based spectral model for isotropic textures. *ACM Transactions on Applied Perception*, 3(4):376–411, 2006.

[11] Andrew Howes and Richard M. Young. The role of cognitive architecture in modeling the user: Soar's learning mechanism. *Human-Computer Interaction*, 12(4):311–343, 1997.

[12] Radu Jianu, Cagatay Demiralp, and David H. Laidlaw. Exploring brain connectivity with two-dimensional neural maps. In *IEEE Visualization 2010 Poster Compendium*, 2010. `http://vis.cs.brown.edu/docs/pdf/g/Jianu-2010-EBC.pdf.html`.

[13] Radu Jianu, Kebing Yu, Vinh Nguyen, Lulu Cao, Arthur Salomon, and David H. Laidlaw. Visual integration of quantitative proteomic data, pathways and protein interactions. *IEEE Trans. on Visualization and Computer Graphics*, September 2009. `http://vis.cs.brown.edu/docs/pdf/g/Jianu-2009-EVI.pdf.html`.

[14] Bonnie John, Alonso Vera, Michael Matessa, Michael Freed, and Roger Remington. Automating CPM-GOMS. In *Proceedings CHI*, 2002.

[15] David E. Kieras and David E. Meyer. An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, 12(4):391–438, 1997.

[16] Jing Li, Jean-Bernard Martens, and Jarke J. van Wijk. A model of symbol size discrimination in scatterplots. In *Proceedings CHI*, 2010.

[17] Zhicheng Liu, Nancy Nersessian, and John Stasko. Distributed cognition as a theoretical framework for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1173–1180, 2008.

[18] Zhicheng Liu and John Stasko. Mental models, visual reasoning and interaction in information visualization: A top-down perspective. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):999–1008, 2010.

[19] Jerry Lohse. A cognitive model for the perception and understanding of graphs. In *Proceedings CHI*, 1991.

[20] Jock Mackinlay. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, 5(2):110–141, 1986.

[21] Jock Mackinlay, Pat Hanrahan, and Chris Stolte. Show Me: Automatic presentation for visual analysis. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1137–1144, 2007.

[22] G. Elisabeta Marai, Joseph Crisco, and David H. Laidlaw. A kinematics-based method for generating cartilage maps and deformations in the multi-articulating wrist joint from ct images. In *Proceedings of the IEEE International Conference of the Engineering in Medicine and Biology Society (EMBS)*, New York, NY, September 2006. `http://vis.cs.brown.edu/docs/pdf/bib/g/Marai-2006-AKB.pdf.html`.

[23] Chris North and Ben Shneiderman. Snap-together visualization: a user interface for coordinating visualizations via relational schemata. In *Proceedings of Advanced Visual Interfaces*, 2000.

[24] Steven Pinker. A theory of graph comprehension. In *Artificial Intelligence and the Future of Testing*. Psychology Press, 1990.

[25] Matthew D. Plumlee and Colin Ware. Zooming versus multiple window interfaces: Cognitive costs of visual comparisons. *ACM Transactions on Computer-Human Interaction*, 13(2):179–209, 2006.

[26] James J. Thomas and Kristin A. Cook, editors. *Illuminating the path: The research and development agenda for visual analytics*. IEEE Computer Society, 2005.

[27] Jarke J. van Wijk. Views on visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):421–433, 2006.

[28] Michelle Q. Wang Baldonado, Allison Woodruff, and Allan Kuchinsky. Guidelines for using multiple views in information visualization. In *Proceedings of Advanced Visual Interfaces*, 2000.

[29] Caroline Ziemkiewicz and Robert Kosara. The shaping of information by visual metaphors. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1269–1276, 2008. `http://cs.brown.edu/people/cziemki/documents/ziemkiewicz08_visual-metaphors.pdf`.

[30] Caroline Ziemkiewicz and Robert Kosara. Laws of attraction: From perceived forces to conceptual similarity. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1009–1016, 2010. `http://cs.brown.edu/people/cziemki/documents/ziemkiewicz10_laws-of-attraction.pdf`.

Mentoring Plan for Postdoctoral Researcher Caroline Ziemkiewicz

Dr. Caroline Ziemkiewicz, as principal investigator, will be involved in managing and coordianting the proposed research as part of the Brown group for an expected appointment of two years. Dr. Laidlaw will act in a mentoring role during this time. Our mentoring plan is designed to aid Dr. Ziemkiewicz in her development of her own long-term research and professional goals.

Interdisciplinary Collaboration. Caroline will work closely with students and faculty in our group to pursue a research agenda that involves close collaboration with scientific researchers, cognitive science, and computer science. This research will include regular collaboration with cognitive scientists and scientific domain experts. The diversity of scientific research labs involved in this work will give Caroline the chance to learn about a broad array of relevant subjects and to apply visualization research in a generalizable way independent of specific data content. The working relationships encouraged by this process will also help to establish interdisciplinary collaborations that can potentially follow Caroline as she continues her career beyond the postdoctoral level.

Publication and Training in Grant Preparation. Caroline will receive assistance from Dr. Laidlaw in the process of writing and publishing these results in reputable journals and conferences. She will be involved in preparing further proposals, as a Co-PI or a PI, to garner funding to support work on subsequent projects, and we will jointly determine how to appropriately share any resulting awards after her departure from Brown.

Career Counseling and Guidance. Dr. Laidlaw will be available to meet with Caroline weekly throughout the period of the mentorship, focusing on her evolving career and research goals. Initially, we will work to define these goals; subsequently we will evaluate whether she is on an appropriate trajectory to reach them. Progress reports at the end of each semester will be used to ensure that this trajectory is being followed and any areas requiring additional guidance. This teleological approach to mentoring naturally supports technical advising on the research itself while preserving the higher-level goals of both the mentee and mentor.

Guidance on Improving Mentoring and Teaching Skills. Caroline will be closely involved in mentoring a graduate student who will be funded under the proposed research, giving her the opportunity to learn how to guide a student towards a research goal. She will also participate in the teaching of several relevant classes taught by Dr. Laidlaw, including the Interdisciplinary Scientific Visualization course and a course titled Virtual Reality Design for Science, which is offered jointly with the Rhode Island School of Design and will provide a unique opportunity to learn how to teach in an interdisciplinary environment. She will also help to advise students performing independent studies. She will also lead group discussions in research meetings and will have the opportunity to teach her own computer science course at Brown. In addition, presenting the findings of this research at conferences will give Caroline the opportunity to practice presentation and lecturing skills and to gain exposure in the visualization research community at large.