THE ANALYSIS AND VISUALIZATION OF ELECTROENCEPHALOGRAPHY DATA USING A PROVENANCE-ENABLED ENVIRONMENT AND ITS APPLICATIONS TO VISUALIZATION

by

Erik W. Anderson

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Computer Science

School of Computing

The University of Utah

December 2011

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Erik W. Anderson**

has been approved by the following supervisory committee members:

Claudio Silva	, Chair	**7/19/2011** Date Approved
Juliana Freire	, Member	**7/19/2011** Date Approved
Guido Gerig	, Member	**7/27/2011** Date Approved
Christopher Johnson	, Member	**7/19/2011** Date Approved
Gilbert Preston	, Member	**8/3/2011** Date Approved

and by **Alan Davis** , Chair of

the Department of **School of Computing**

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

Electroencephalography (EEG) remains a common data collection technique in the field of Neuroscience. Improvements in EEG devices are often expressed as increases in temporal or spatial resolution. The concrete result of these improvements is an ever-increasing amount of data to be processed. Appropriately analyzing EEG data is a complex process requiring detailed provenance to be recorded at each step of the processing workflow. VisTrails is a provenance-based workflow system adapted to enable the processing and visualization of EEG data. To create a suite of tools amenable to EEG analysis, VisTrails was extended to include functionality for manipulating arrays and matrices, digital signal processing, and conversion utilities for visualization. The resulting tool has proven useful in several studies rooted in neuroscience and has led to several new insights about working memory and cognition.

Working memory has been described as short-term retention of information that is no longer accessible in the environment, and the manipulation of this information for subsequent use in guiding behavior. Working memory is viewed as a cognitive process underlying higher-order cognitive functions. Studies show psychomotor processing speed and accuracy account for considerable variance in neural efficiency. This study compared the relative effects of active and sham 10 Hz repetitive transcranial magnetic stimulation applied to dorsolateral prefrontal cortex on indices of neural efficiency in healthy participants performing a working memory paradigm that models the association between working memory load and task behavior. Previous studies identified a relationship between diminished neural efficiency and impaired working memory across a broad array of clinical disorders. In the present study, the authors predicted there would be a main effect of stimulation group on accuracy and processing speed, hence, neural efficiency. We observed a main effect of stimulation for reaction time without an effect on accuracy; even so, there was a robust effect of stimulation on neural efficiency.

Effectively evaluating visualization techniques is a difficult task often assessed through feedback from user studies and expert evaluations. We present an alternative approach to visualization evaluation in which brain activity is passively recorded using electroencephalography (EEG). These measurements are used to compare different visualization techniques in terms of the burden they place on a viewer's cognitive resources. In this work, EEG signals and response times are recorded while users interpret different representations of data distributions. This information is processed

to provide insight into the cognitive load imposed on the viewer. This work describes the design of the user study performed, the extraction of cognitive load measures from EEG data, and how those measures are used to quantitatively evaluate the effectiveness of visualizations.

To my family.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

CHAPTER 1

INTRODUCTION

1.1　Motivation

In the art community, provenance refers to the source of an object, its history of owners and caretakers, the preservation and restoration techniques applied to it, and the various documents accompanying sales or transferral of the artefact. The idea of provenance also translates to scientific experimentation in the form of detailed notes regarding the particulars of an experiment, the computations performed, and the results and insights gained. Recording provenance in science helps to interpret, analyze, and reproduce the results of an experiment, calculation or simulation.

Provenance is not new to scientific endeavours. Laboratory notebooks have traditionally been the recording medium of choice for scientific provenance, but with the advent of high performance computing, the volume of data quickly overwhelms this method. Fortunately, recent efforts have been focused on this problem [21, 143].

The VisTrails System is a provenance management system designed to support exploratory interaction with both computational and visualization tasks [52]. Its extensible architecture and support for advanced visualization and computational functionality make it amenable for enabling rich provenance capture in EEG data analysis and visualization.

Electroencephalography (EEG) allows scientists to study the awake and alert brain in a noninvasive way by measuring voltage differences across the scalp. These voltage differences are generated by the coordinated electrical activity of collections of neurons working together in the cerebral cortex. By examining the expression of these voltages at the scalp, neuroscientists gain additional understanding of the macroscopic function of the brain in terms of its functional units.

Although there are other methods to examine the alert human brain, EEG has several distinct features that are particularly important to its use in Neuroscience studies. Since EEG is measuring voltages expressed on the scalp and generated by the neural substrate performing work, generating a view of coordinated activity in the brain is possible. Additionally, because raw EEG data are represented as real valued timeseries collected at the sensor locations, the power of digital signal processing is leveraged to provide an effective suite of tools during analysis. Finally, EEG sensors

are often anchored in a flexible headset. EEG headsets are then connected, sometimes wirelessly, to a signal receiver allowing both movement and change in posture and orientation.

EEG is an important tool for the study of the brain, but its strengths also impose unique challenges that must be addressed. High resolution EEG data collected at many spatial locations may provide a low level representation of brain activity, but making sense of these data remains difficult. Visualization methods have been devised to provide scientists with spatial contexts upon which hypotheses may be formed, analyzed, and tested [82, 150].

Working memory has been described as short term retention of information that is no longer accessible in the environment, and the manipulation of this information for subsequent use in guiding behavior [43]. Working memory is widely viewed as a cognitive process that underlies an array of higher order cognitive functions, such as reasoning [129], planning [62], and problem solving [44], for instance. Over the past few decades, such studies have developed a substantial body of convergent evidence that supports a critical role for prefrontal cortex (PFC) in mediating working memory performance. For example, in a study aimed to identify the relationship between working memory demand and task performance, Rypma et al. found that task factors exert their effects largely during working memory encoding [129]. Subject factors influence on task performance occurred mainly during working memory retrieval. In an fMRI study, subjects performed a working memory task that required them to maintain from one to eight letters over a brief delay [127]. Neural activity was measured during encoding, maintenance, and retrieval task phases. With increasing memory load, the researchers observed that reaction time increased and accuracy decreased. Additionally, a decline in ventrolateral PFC (VLPFC) activation was seen during encoding, whereas dorsolateral PFC (DLPFC) activation increased during maintenance and retrieval. The authors concluded that VLPFC mediates working memory storage and that DLPFC mediates memory organization processes that facilitate supracapacity working memory storage. Interestingly, high performing subjects in toto showed less activation than low performing subjects, hence, suggesting that high performers utilized fewer neural resources in resolving task requirements. In comparison to low performers, such task performance could plausibly be seen as an example of greater neural efficiency. Even so, higher scoring performers showed activation increases with increasing memory load during maintenance and retrieval in lateral PFC. Overall, low performing subjects showed more activation than high performing subjects, but high performers showed minimal activation increases in DLPFC with increasing memory load. The authors suggested that their results indicate that individual differences in neural efficiency and cognitive strategy mediate individual differences in working memory performance [127] (Rypma et al., 2002).

These studies, taken along with the work of Wolfgang Klimesch show that working memory performance is directly measurable by careful inspection of EEG [77, 78]. Direct inspection of working memory performance enables scientists to develop mechanisms to manipulate the neural rhythms subserving the working memory systems. One such way of manipulating memory is with repetitive transcranial magnetic stimulation (rTMS) [11, 35, 78, 135]. rTMS delivers a series of magnetic pulses to the cortex where it then induces changes in the electrical potentials in the collection of neurons in that area. In this manner, the function of neural assemblies is changed in a localized area and in a controlled way.

By manipulating the rhythms and oscillations in cortex, the systems that are controlled by neural assemblies may be influenced. Evidence to support the hypothesis that rTMS manipulates the performance of neural assemblies, particularly those associated with working memory, has been generated by several studies [7, 94, 95, 120]. Using rTMS to manipulate the neural circuit subserving working memory may affect neuropsychiatric disorders with working memory or cognitive dysfunction. In this way, new treatments may be developed that directly influence brain function without the use of pharmaceuticals. One neuropsychiatric disorder that may be amenable to this technique is schizophrenia.

Schizophrenia has been recognized for more than a century as a psychotic disorder in which psychosocial recovery is uncommon, principally due not to psychotic symptoms which are relatively easy to treat, but to disabling impairment of cognitive functions first recognized by Emil Kraepelin in 1893 [87]. Kraepelin described poor outcomes in social functioning, associated with impairments of attention, motivation, problem solving, learning, and memory as the principal features of this disorder. Kreapelin's assessment still holds — less than 15% of patients with schizophrenia recover [27, 99]. There is still no treatment for cognitive dysfunction seen in schizophrenia or other neuropsychiatric disorders [67, 102, 123]. Recent reports from the NIMH supported MATRICS researchers regarding the development of a consensus for a standardized battery of neuropsychological instruments to identify and measure the principal cognitive deficits of schizophrenia reinforces the importance of cognitive impairment in schizophrenia [64].

Although working memory is able to be manipulated through the careful application of rTMS, the ability to monitor its performance also presents new and exciting opportunities. Since working memory and cognitive processes are linked [14], the mechanisms developed to monitor these systems may be able to give us new insights into the cognitive loads imposed on people by certain visualization techniques. By reducing the cognitive demands of interpreting data, new methods may be developed to more quickly and accurately portray the most salient aspects of scientific data to scientists.

Efficient visualizations facilitate the understanding of data sets through an appropriate choice of visual metaphor. Within the field of visualization, there exist numerous display strategies, many of which can be applied to similar types of data. These various techniques often create distinct imagery, emphasizing particular data characteristics or visualization goals. In most cases, several rendering techniques are appropriate; however, some methods may present salient information more quickly and accurately. The choice of best visualization technique for a particular data set is difficult to make. The visualization expert must not only determine an appropriate technique for the type of data, but also ensure the chosen method will answer the questions posed by domain experts. The difficulty of this choice is exacerbated by the lack of exhaustive visualization evaluation detailing the effectiveness of methods for particular types of inquiry.

Often, evaluation of visualization techniques is conducted through expert assessments and user studies, which typically judge a visualization using verbal feedback and user performance. While some measures of usability and effectiveness are relatively easy to quantify, such as increases in users' response speed or decreases in their error rates, others are problematic. For example, it is difficult to assess improved understanding and insight because those metrics tend to be highly subjective. Approaches to evaluation which rely on verbal feedback can be influenced by personal preference, user expectations, cultural biases within scientific fields, and resistance to change. The work described in this paper strives to evaluate visualization techniques objectively by using passive, noninvasive monitoring devices to measure the burden placed on a user's cognitive resources.

The study we present in this paper explores the amount of work, defined by cognitive load, needed to interpret a visualization. We evaluate some simple visualization methods by measuring the brain activity through electroencephalography (EEG). A framework is defined for the processing and analysis of the acquired EEG sensor data which allows for the interpretation of difficulty of a visualization task. We believe the results of this study to be an important advancement of objective visualization evaluation.

1.2   Contributions

This dissertation presents a suite of processing tools to process, analyze and visualize EEG data, and two studies performed using this provenance-enabled environment. The first study investigates the measurement and manipulation of working memory performance using EEG. The second study uses working memory performace measured by EEG in concert with Cognitive Load Theory to develop a measure of cognitive load associated with various types of visualization.

The main contributions of this dissertation can be divided into three distinct categories: the use of a provenance aware environment for EEG processing, the use of EEG data during the

measurement and manipulation of human working memory, and the use of EEG in user studies of visualization.

The provenance management system, VisTrails, was chosen as the platform of choice for extension to include a suite of tools useful for the analysis and visualization of EEG data. The flexibility and feature set of VisTrails, combined with the large number of preexisting libraries available to it via Python, made VisTrails the best choice for modification [8]. The contributions related to the use of a provenance aware environment are discussed in detail in Chapter 3 and are summarized as follows:

- The extension of the VisTrails Provenance Management System to incorporate analysis and visualization of EEG Data.

- The discussion of the utility of VisTrails in its role processing data collected during neuroscience experiments.

The development and acceptance of reprititive transcranial magnetic stimulation by the neuroscience communities opened new avenues for innovative research. Exposing the brain to this type of stimulation temporarily alters the way in which neuron assemblies communicate, thus manipulating the higher level functions for which they are responsible. A study was performed and is discussed in Chapter 4 by which the extent of manipulation of working memory is found. Additionally, the processing of EEG data to measure the spectral consequences of magnetic stimulation is developed and refined. The contributions related to the use of EEG data to the measurement and manipulation of working memory are as follows:

- A multidisciplinary approach to the collection, analysis, and visualization of EEG data in neuroscience experiments.

- The use of repetitive transcranial magnetic stimulation (rTMS) to induce robust enhancement to working memory performance.

- The spectral analysis of EEG data to extract measures of working memory performance.

Measuring working memory performance provides a window into the function of the brain as the decision making process is carried out. Using the established links between working memory and cognition [14], a user study was developed to explore the cognitive implications of different visualization methods (see Chapter 5). The contributions of the execution and analysis of this study are summarized as:

- The use of EEG to inspect brain activity while interpreting visualizations.

- The use of cognitive load as a objective measure of visualization effectiveness

- The formulation of cognitive load based on its spatial, spectral, and temporal organization.

- The use of working memory as an estimation of cognitive load.

The work presented in this dissertation has been previously been published [6–9, 46, 120].

CHAPTER 2

BACKGROUND

This chapter provides the background for the remainder of this dissertation by reviewing relevant previous works. Section 2.1 reviews provenance and provenance management systems relevant for the discussions in Chapter 3. Section 2.2 presents background on electroencephalography (EEG) and its significance in assessing working memory performace necessary for the work described in Chapter 4. Finally, Section 2.3 highlights previous work exploring Cognitive Load Theory and evaluation methods for visualization. This background relates to the discussion presented in Chapter 5.

2.1 Provenance

The word "provenance" is commonplace when studying art or archaeological artifacts. The provenance of an object captures the entire chain of custody from the time it is created to the present, an important piece of information in art history and achaeology. Provenance is not new to scientific endeavours. Laboratory notebooks have traditionally been the recording medium of choice for scientific provenance, but with the advent of high performance computing, the volume of data quickly overwhelms this method. Fortunately, recent efforts have been focused on this problem [21, 143].

2.1.1 Digital Provenance for Scientific Data

There are many ways to represent provenance in a computational environment. Freire et al. provide a full discussion on the various types of provenance and the mechanisms by which it is captured, stored, and queried [52]. Provenance is expressed in two distinct forms — *prospective provenance* and *retrospective provenance*. Prospective provenance records the specification of a task whereas retrospective provenance captures the operations performed on data and information about the execution environment to ensure a proper history [40]. Prospective provenance may be thought of as a recipe to generate a data product while retrospective provenance is similar to a detailed log of a computational task.

Several systems for managing provenance in scientific computing and visualization are available. Many of these systems use a prospective provenance pardigm represented as a workflow to de-

fine the operations performed on the input data. Environments such as Kepler [22] and Taverna [73] have created a provenance management layer to support their workflow system. However, systems such as the VisTrails Project have been designed around the capture of provenance information rather than adding functionality to an already existing system [1].

2.1.2   The VisTrails Environment

The VisTrails System is a provenance management system designed to support exploratory interaction with both computational and visualization tasks [52]. Its extensible architecture and support for advanced visualization and computational functionality make it amenable for enabling rich provenance capture in EEG data analysis and visualization.

The design of the VisTrails system places importance on the collection and maintenance of provenance [84]. VisTrails' visual programming model, easy extensibility, and rich feature set allowed for rapid implementation of the various components necessary to support EEG processing and analysis.

Three primary views in VisTrails enable the interaction with provenance information, workflow components, and final visualization products. Change based provenance regarding the contruction of various workflows is captured and displayed in the *Version Tree*. Each oval in the Verstion Tree represents a complete workflow. Selecting a specific version allows interaction with its workflow via the *Workflow View*. In many cases, the results of workflow execution are displayed in the *Spreadsheet Window*. Each cell of this spreadsheet is capable of rendering the results as complete 3D visualizations, 2D graphs, or even annotated tables.

VisTrails enables the creation and execution of workflows composed of computational elements called *modules*. Modules are distributed in *packages* and connected to one another to generate an executable workflow. Although many workflows use a small collection of static operations, VisTrails provides some modules designed to increase its flexibility. In particular, the Python Source module allows a user to write custom code to interact with the workflow in nonstandard ways. Vis-Trails is also capable of interacting with other execution engines. By interfacing with other Python packages such as mlabwrap (`http://mlabwrap.sourceforge.net`), VisTrails gains the ability to run algorithms in the Matlab environment. In this way, VisTrails may utilize the algorithms implemented in the EEGLab suite of tools.

2.2   Electroencephalography and Working Memory

Electroencephalography (EEG) measures brain activity by monitoring electrical potential generated by the cortex on the surface of the scalp. Each voltage sensor in an EEG system measures

activity at a discrete location and expresses the information in the form of a time series. Neuroscientists have utilized EEG to gain a richer understanding of brain functions [28, 58].

2.2.1 EEG Processing and Visualization

EEG analysis utilizes various techniques from the signal processing community. One problem of considerable interest during EEG analysis is signal mixing as multiple signal sources in the brain are recorded by each EEG electrode. Independent Component Analysis (ICA) [75] provides a method for separating independent signals from the EEG mixture [76] recorded during Event Related Potential studies.

Raw data from a single EEG sensor are expressed as a time series. This representation of data is well suited to a variety of processing methods [100, 107]. Additionally, spectral decomposition methods play an important role in the analysis of EEG data [7, 78, 120]. Unfortunately, there is no consensus on the specific transforms to use during analysis. Windowed Fourier Transforms [15], Wavelet Transforms [137], Gabor Transforms [144], and S-Transforms [8] have all been applied to EEG data.

The variety of techniques used to process EEG data presents a problem for the development of unified analysis environments. Two such environments, EEGLab [41] and FieldTrip [109], extend the Matlab Environment (`http://www.mathworks.com`), implementing both low level and high level functionality to provide EEG specific features. Extending a computational environment is advantageous as it enables the use of existing, well tested components.

Many different techniques have been developed for visualizing time series data. Aigner et al. present a survey covering visualization methods spanning a number of temporally sensitive datasets [3]. One of the primary goals of visualizing time series data is to extract patterns, or cycles, within the ensemble. VisTree [92] organizes collections of time series to visually detect patterns in them. Similarly, Yankov et al. use multiple dot plots to discover patterns in a spreadsheet like array [164].

There have also been several visualization methods taking into account the unique challenges related to EEG data. Using the known spatial layout of the EEG sensor net enables additional cues to inform the scientists. Spectral Topomaps [82], Tiled Parallel Coordinate Maps [151], and Functional Unit Maps [150] all leverage this additional information to enhance the visualization. Using spatial information in these visualizations helps scientists gain additional insight into brain activity.

2.2.2 Working Memory

Working memory is responsible for the retrieval, manipulation, and processing of task related information and has functional importance to a variety of cognitive activities including learning, reasoning, and comprehension [14]. It is often useful to think of the working memory system in terms of a computer architecture in which working memory acts as the central processing unit (CPU) with direct connections to temporary data buffers (RAM) in the form of short term memory, and external communications (IO) through sensory perceptions and resulting reactions [14]. Of course, the actual working memory system is much more complex than a computer, and therefore dividing up the processes of the system is not always possible, as many of the functions occur across the same neural substrate [37].

Through the use of EEG, neuroscientists have learned that working memory is organized in both the spatial [13, 37] and spectral [45, 78] domains. Although a strict spatial segmentation of the brain in terms of working memory activity is impossible, Braver et al. show that the working memory processing is measurable in the prefrontal cortex of the brain [25] while Constantinidis et al. explore a more complete neural circuit for spatial working memory [38].

2.3   Cognition and Visualization

2.3.1   Cognitive Load Theory

Cognitive load theory [148] describes the relationship between the capacity of working memory and the cognitive demands of a particular task. The core of the theory is that people have a limited cognitive capacity during learning and problem solving tasks. The way in which information is presented can affect the amount of load placed on the working memory system and thus affect performance [47]. Cognitive load theory distinguishes three types of cognitive load: *germane, intrinsic, and extraneous* [33]; each distinctly affecting learning and decision making. The combination of the three types characterize the overall cognitive load [140].

2.3.1.1   Germane cognitive load

Germane cognitive load is the load devoted to learning new cognitive schema [148]. These schema are internal representations formed in the learning process which are used over and over and may be relevant to many tasks. Once these cognitive schema are in place, the contribution of germane cognitive load to the overall load is minimal.

2.3.1.2   Intrinsic cognitive load

Intrinsic cognitive load describes the demands on working memory capacity generated by the innate complexity of the information being examined [148]. This load represents the portion of

overall cognitive load that is influenced by the difficulty of the underlying task at hand and cannot be manipulated by the design of the task. An example of intrinsic cognitive load is the inherent challenge involved in adding two numbers compared to the greater challenge in solving more advanced arithmetic problems.

2.3.1.3 Extraneous cognitive load

Extraneous cognitive load measures the additional load placed on users by the design of a task [114]. This type of load can be controlled by the way information is presented [140]. For example, a table of numbers describing 1-dimensional distribution data imposes more extraneous cognitive load on a user to understand than the visual representation of a histogram containing the same data.

2.3.2 Visualization Evaluation

A substantial barrier to the evaluation of visualization techniques is the complexity of the task. Not only must a technique appropriately portray the data, but it also must sufficiently outperform equivalent rendering techniques. While appropriate measures for these requirements are difficult to formulate, there exists the additional challenge that most visualization problems are highly application dependent; visualization techniques that are validated as effective for one particular type of problem may not perform well for another one, even if the two are similar. Many visualization techniques are presented with evaluations which rely on technical improvements such as speedups, or the management of larger data sets. However, the use of human factors, user studies or expert evaluations is becoming more common.

User studies are effective ways of evaluating everything from visualization methods [88, 133] to complex environments such as airplane cockpits [134] and surgical simulators [121]. These classes of user studies generally use post experiment surveys in conjunction with timing and task related data to form a foundation for additional statistical analysis. These user studies leverage both empirical data collected during the user task as well as subjective data collected after the experiment.

While user studies have become an important tool in the assessment of visualization methods, they are not always the best evaluation technique. Kosara, et al. [85] show that user studies are effective at answering specific questions, such as "Does a specific method of streamline rendering show vorticities better than others?" Similarly, Cleveland and McGill [36] use evaluation studies to answer focused questions about data visualized in different ways.

Human factors play an important role in the study of the impact of scientific visualization on research. They are particularly important during the evaluation of visualization systems. An example of this type of system is Kosara, et al. using semantic depth of field [86] in which renderings

strive to induce perceptual changes in the user. Tory and Möller [153] offer a thorough discussion of human factors in both user study methods, and also in visualization design.

CHAPTER 3

VISTRAILS: EEG VISUALIZATION AND ANALYSIS IN A PROVENANCE RICH ENVIRONMENT

This chapter is organized as follows: in Section 3.1 we discuss the VisTrails System and how it is extended to provide the necessary functionality. Section 3.2 tackles the problem of realtime collaboration. Section 3.3 discusses the specific singal processing and visualization features added to VisTrails. Section 3.4 then gives presents a concrete example of processing and visualization within the VisTrails environment while Section 3.5 provides a discussion of the work.

3.1 The VisTrails System

Figure 3.1 provides an illustration of the VisTrails environment. Three primary views in Vis-Trails enable the interaction with provenance information, workflow components, and final visu-alization products. Change based provenance regarding the contruction of various workflows is captured and displayed in the *Version Tree*. Each oval in the version tree represents a complete workflow. Selecting a specific version allows interaction with its workflow via the *Workflow View*. In many cases, the results of workflow execution are displayed in the *Spreadsheet Window*. Each cell of this spreadsheet is capable of rendering the results as complete 3D visualizations, 2D graphs, or even annotated tables.

VisTrails enables the creation and execution of workflows composed of computational elements called *modules*. Modules are distributed in *packages* and connected to one another to generate an executable workflow. Figure 3.2 depicts the directed graph that performs a Fast Fourier Transform (FFT) on an input EEG and displays the results. Although this workflow uses a small collection of static operations, VisTrails provides some modules designed to increase its flexibility. In particular, the Python Source module allows a user to write custom code to interact with the workflow in nonstandard ways. VisTrails is also capable of interacting with other execution engines. By interfac-ing with other Python packages such as mlabwrap (`http://mlabwrap.sourceforge.net`), VisTrails gains the ability to run algorithms in the Matlab environment. In this way, VisTrails may utilize the algorithms implemented in the EEGLab suite of tools.
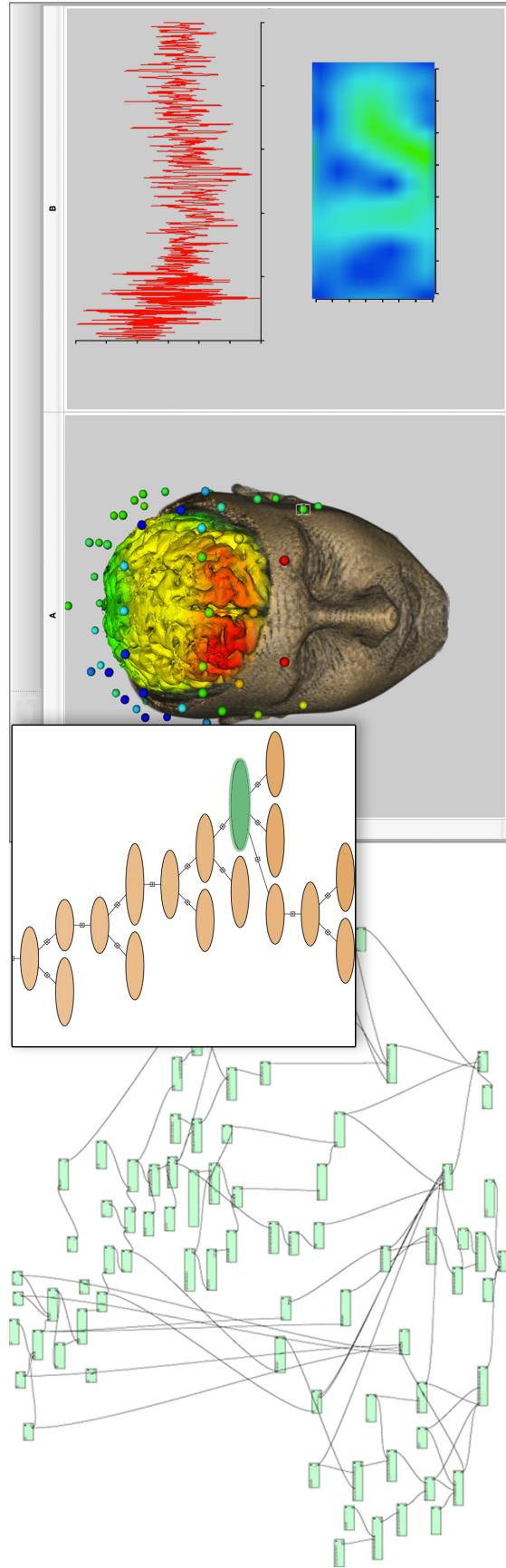
Figure 3.1: The VisTrails Interface: The workflow view (left) allows the user to create a workflow for execution. The visualizations formed by workflows may be rendered on the VisTrails Spreadsheet (right). Each of the workflows in VisTrails is contained in a unique entry in the VisTrails Version Tree (middle). In this case, only part of the tree is shown and the workflow and visualization are defined in the version highlighted in green.
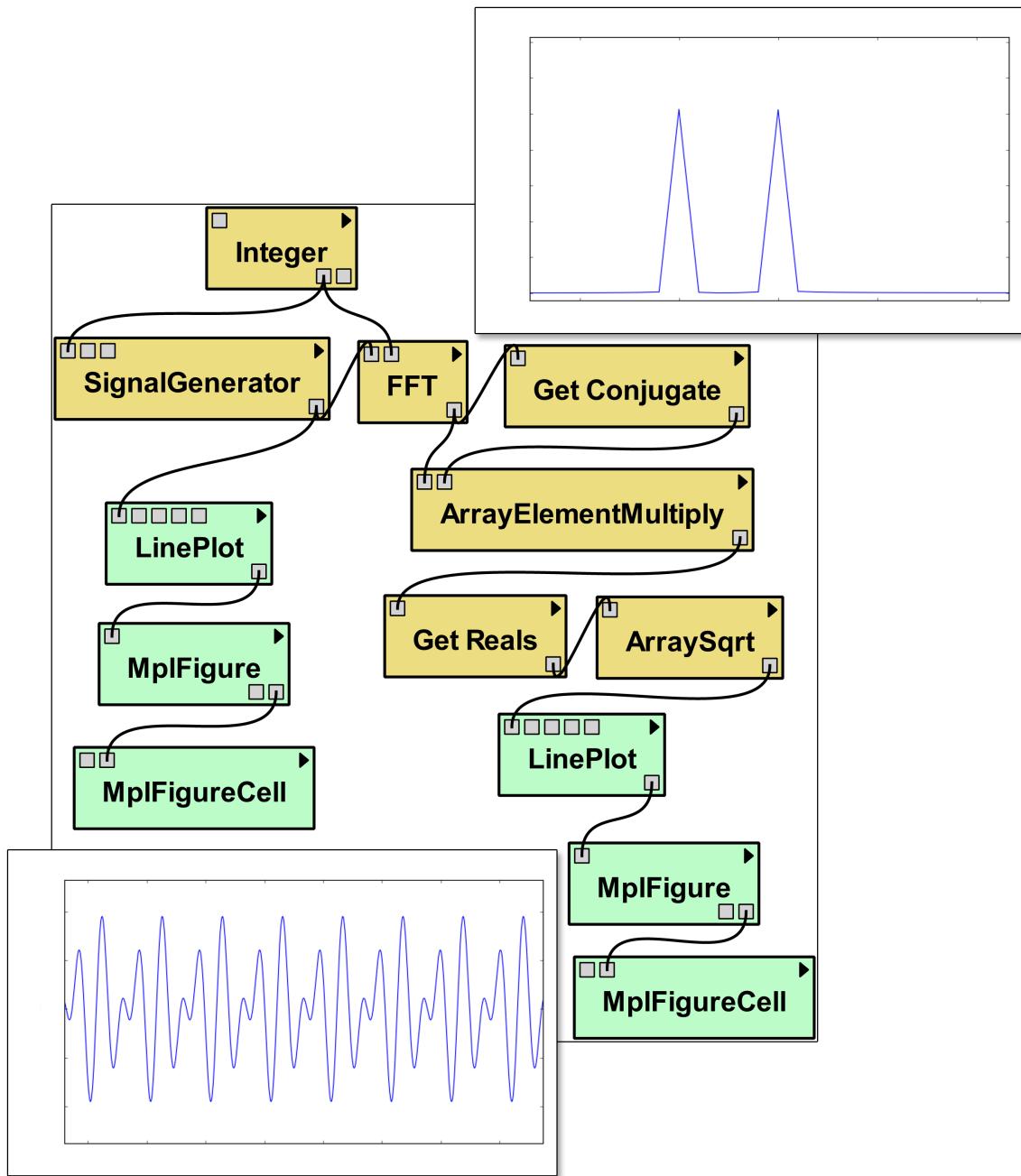
Figure 3.2: A VisTrails workflow that generates and displays a signal composed of two frequencies. The workflow is able to display both the original signal (bottom left) as well as its Fourier Transform (top right) from within the same workflow.

The VisTrails system has proven to be successful in a variety of applications including teaching visualization [142], combining simulation and visualization with provenance [72], collaboratively designing analyses and visualizations [46], and in comparative visualization [9]. The ability of VisTrails to function not only as an analysis and visualization platform, but also as a teaching tool is important for multidisciplinary collaboration. By enhancing communication between neuroscience experts and visualization experts, VisTrails enables deeper and more robust collaboration between different scientific fields.

3.1.1   Provenance-Enabled Operations

One of the primary reasons VisTrails is a good candidate for EEG processing is its provenance management system. The VisTrails version tree (See Figure 3.1) not only manages the evolution of a collection of workflows, but it also provides advanced interaction with the recorded provenance and other metadata. Users may issue text based queries to find specific nodes on the version tree containing the query text in the node name, notes, or even which user was responsible for the change. However, utilizing provenance enables additional types of queries to be issued and operations to be performed [29, 53, 136].

One way that provenance information is leveraged in VisTrails to enhance user experience is through the use of *subworkflows*. A subworkflow functions as a single module within an existing workflow, but is composed of multiple modules and connections reflecting its desired functionality. Because a subworkflow is constructed as if it were simply a special type of VisTrails workflow, the provenance associated with it is maintained allowing changes in its definition to be rapidly applied to all workflows using it. By creating collections of subworkflows, complex computations may be expressed in workflows as single modules similarly to the transparent abstraction afforded by functional programming [84].

Searching for specific workflows is often difficult, particularly when the search space is large. Additionally, constructing a query to find a specific set of workflows is challenging if the defining features of the target workflows are based in their structures. VisTrails solves this problem in its *query by example* interactions [53]. Query by example allows users to search for specific workflows by constructing interesting workflow fragments to search for. An example of this type of structural query is seen in Figure 3.3. Allowing users to contruct queries as if they were manipulating workflows provides a more intuitive search mechanism to find workflows of interest.

Provenance is also useful for quickly determining the differences between two different workflows. To that end, VisTrails employs a *visual diff* by which the provenance metadata are analyzed to determine the structural differences underlying functional changes [53]. Figure 3.4 illustrates the use of a visual diff to highlight the structural changes required to change rendering methods.
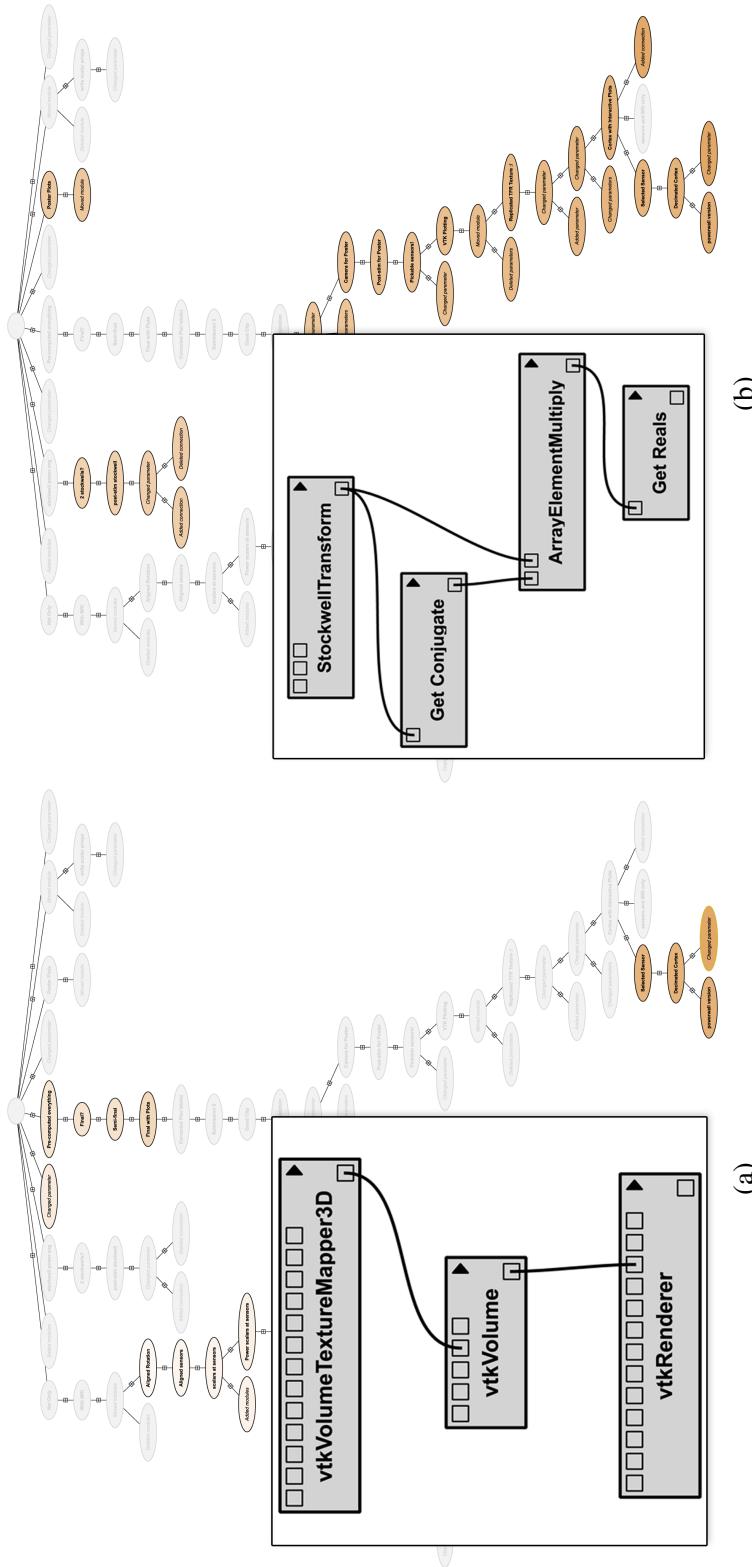
Figure 3.3: Structural queries in VisTrails are performed through the *query by example* interface. Constructing partial workflows defines the query. Query results are reflected in the version tree. A query searching for the modules responsible for texture based volume rendering (a) has a different set of results from a query concerned with time frequency decomposition (b). It is possible for workflows to contain elements of both queries, as is reflected in the highlighting of common versions.

Because the VisTrails provenance mechansims record changes to workflows, a visual diff is much like an arithmetic difference. Scheidegger et al. explore the idea of using such a difference to automatically update workflows in VisTrails [136]. This method of *visualization by analogy* allows workflows to be easily updated to reflect changes in visualization, processing, or analysis. This process is detailed in Figure 3.5.

3.2 Collaboration in VisTrails

In order to support realtime collaborative design workflows, we need a provenance architecture that supports a collection of versioned workflows and a centralized provenance repository that all collaborators can access. We require a versioning system because each user needs to know how their collaborators' work relates to their own. More importantly, we need to protect the users' work; we should not blindly erase or update their own changes. A centralized repository is needed to manage all the workflows and to provide the means for notifying collaborators when changes occur. The combination of these two methods not only allows users to efficiently share collections of workflows, but also enables them to see the entire history of the workflow specifications as they develop in realtime, regardless of how many users collaborate on the project.

3.2.1 Workflow Evolution Provenance

Because we expect to encounter a large number of changes to a workflow specification, especially in a collaborative environment, it can be inefficient to store specifications for all different versions of the workflows. The change based provenance model [54] provides a concise representation for workflow evolution history. This model captures the changes applied to a series of workflows, akin to a database transaction log. As a user modifies a workflow (e.g., by adding a module, changing a parameter or deleting a connection), the provenance mechanism transparently records each change action. We can then reconstruct any workflow specification by replaying the sequence of captured changes from an empty specification to the desired version.

The change based model not only captures changes as a workflow evolves, but it also presents external changes to collaborators in a meaningful way. An important feature of this representation is that it can be visualized as a *version tree*, where each node corresponds to a workflow specification and each edge corresponds to the sequence of changes that transforms the parent specification into the child. Because the version tree captures *all* changes, users have great flexibility for exploring different alternatives without worrying about losing the ability to go back to a specific version. They can perform arbitrary undos and redos—any workflow version is easily recalled by selecting the corresponding node in the version tree. Additionally, users can easily see how their collaborators have taken different approaches to solving related problems and how their techniques relate to their
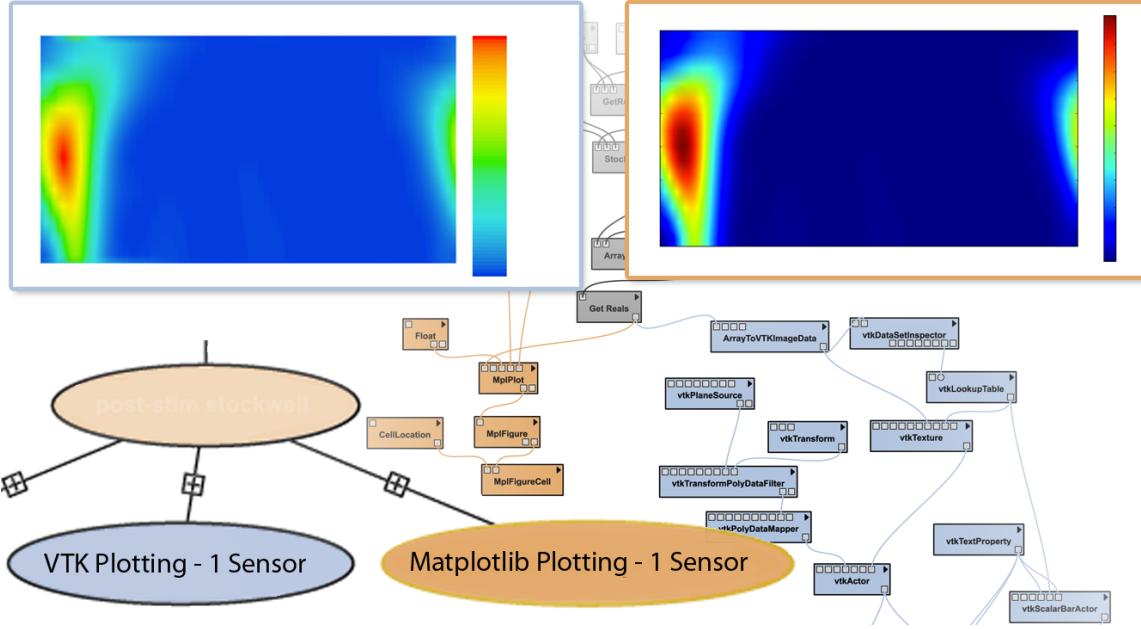
Figure 3.4: The visual diff in VisTrails highlights the structural and functional differences between two individual workflows. In this example, modules highlighted in blue are unique to the workflow using VTK to plot 2D spectral data. Similarly, modules in orange are responsible for plotting the same data using MatplotLib.

own ideas. As discussed below, we leverage this layout to inform users of changes without forcing them to immediately consider or integrate those changes.

3.2.2   Centralized repository

In order to efficiently capture and broadcast workflow changes, we use a relational database management system (RDBMS) for our centralized repository. We chose to use an RDBMS because these systems provide secure access protocols, support concurrent transactions from multiple users, and include trigger mechanisms for alerting users when the database is updated. These features are essential to ensure data consistency and to support realtime updates in our collaborative infrastructure. Other kinds of database systems that support these features could also be used in our infrastructure.

To use an RDBMS for our repository, we need to map the necessary provenance information to a relational schema. Because we use the change based representation, a collection of related workflows is stored as a tree. This tree contains metadata and an ordered set of actions that correspond to user modifications to workflows. Each action, in turn, consists of a sequence of atomic operations. For example, a paste *action* that adds a set of modules and connections to an existing workflow contains a sequence of *operations*: `add module, add connection,`
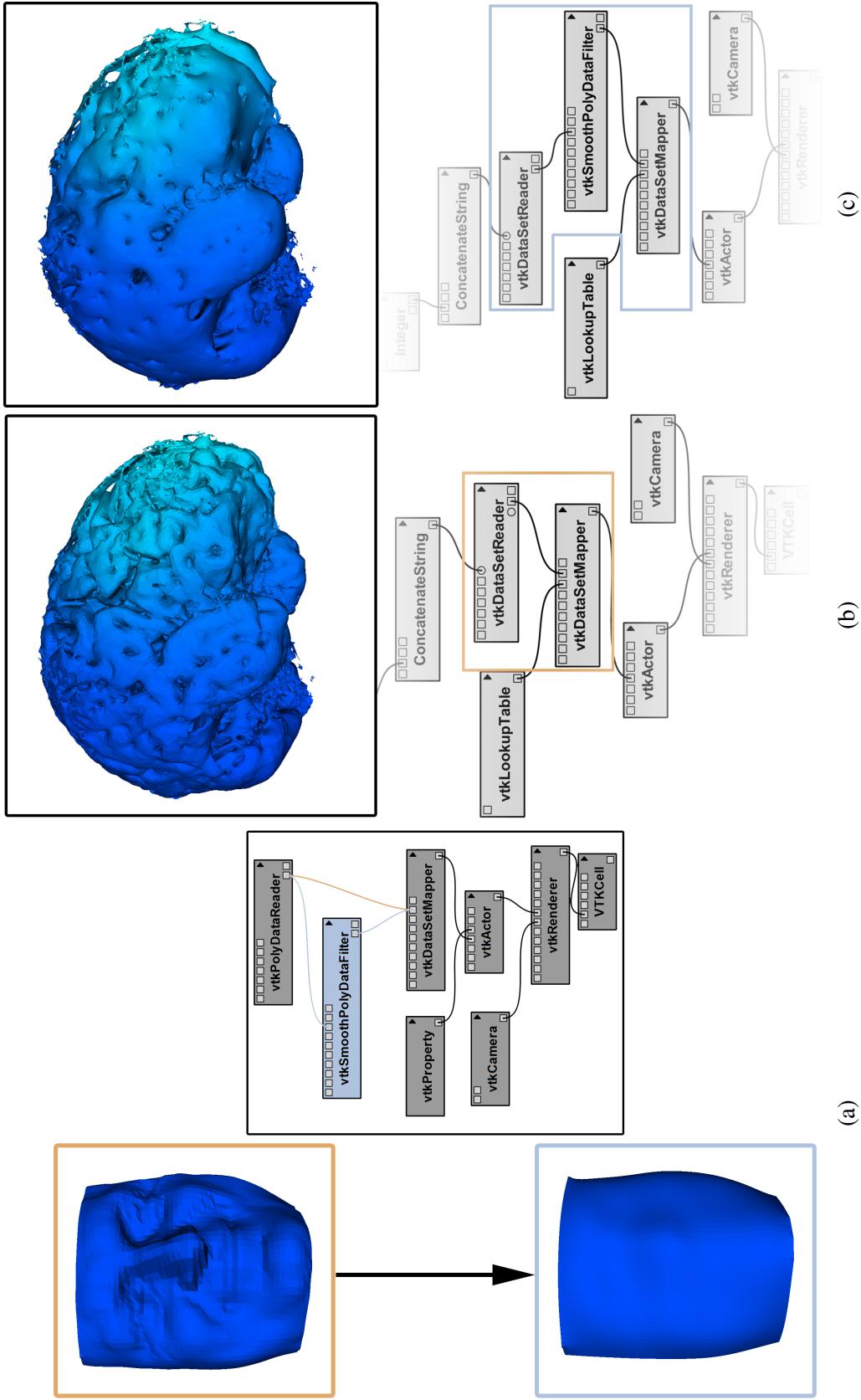
Figure 3.5: Visualization by analogy in *VisTrails* enables the automatic modification of existing workflows. The Visual Diff functionality of VisTrails is used to create an analogy template (a). This template is then applied to a target workflow (b) to automatically apply the relevant changes. The result of this operation is a new workflow (c). Because the analogy inspects more than just module names when it is applied, appropriate workflows are created even if the upstream and downstream modules are mismatched from template to target.

etc.. An operation, besides its data payload (e.g., module specification, connection specification, parameter value), includes metadata (e.g., the user who performed the action and annotations). Each of these entities (actions, operations, payloads) is stored in its own table, permitting a normalized (redundancy free) representation. In addition to storing the changed based representation of workflow evolution, the schema also supports explicit workflow specifications and workflow execution information. Execution information can be important when users are unfamiliar with the collection of workflows and wish to know which workflows are routinely used and which workflows were successfully executed.

3.2.3   Synchronized Design

One of the contributions of this paper is a new method for automatically capturing workflow changes performed by multiple users and alerting them about these changes immediately and unobtrusively. This allows users, in different geographically distributed locations, to collaboratively design and refine workflows, like in the scenario illustrated in Figure 3.6. We accomplish this by committing the local changes (performed by each individual user) to a centralized repository, sending the changes out from the repository to each collaborator, and adding the changes to each collaborator's local version tree. Note that we are not merging workflow specifications but synchronizing workflow evolution provenance. Each collaborator can continue their work and they need not even view the new changes. Before describing the implementation of our prototype, we describe the algorithm for synchronizing the version tree.

3.2.4   Algorithm

There are two key requirements for our algorithm. First, we need a way to save data from a local version tree to the centralized repository. Second, we need a way to load data from that repository to update the collaborators' local version trees. Below, we describe the mechanisms we developed to satisfy these requirements.

Recall that the version tree is induced by a set of actions A. Each action $a \in A$ has a unique identifier derived by the function $\textbf{id} : A \to \mathbb{N}$, where $\textbf{id}$ assigns the smallest unassigned integer to a new action. This function is trivially monotonic: given $a_1, a_2 \in A$,

$$\textbf{id}(a_1) < \textbf{id}(a_2) \iff a_1 \text{ was added before } a_2$$

We will leverage this property to easily determine what has changed in a given version tree. Specifically, let

$$N(A) = \max_{a_i \in V} \textbf{id}(a_i)$$
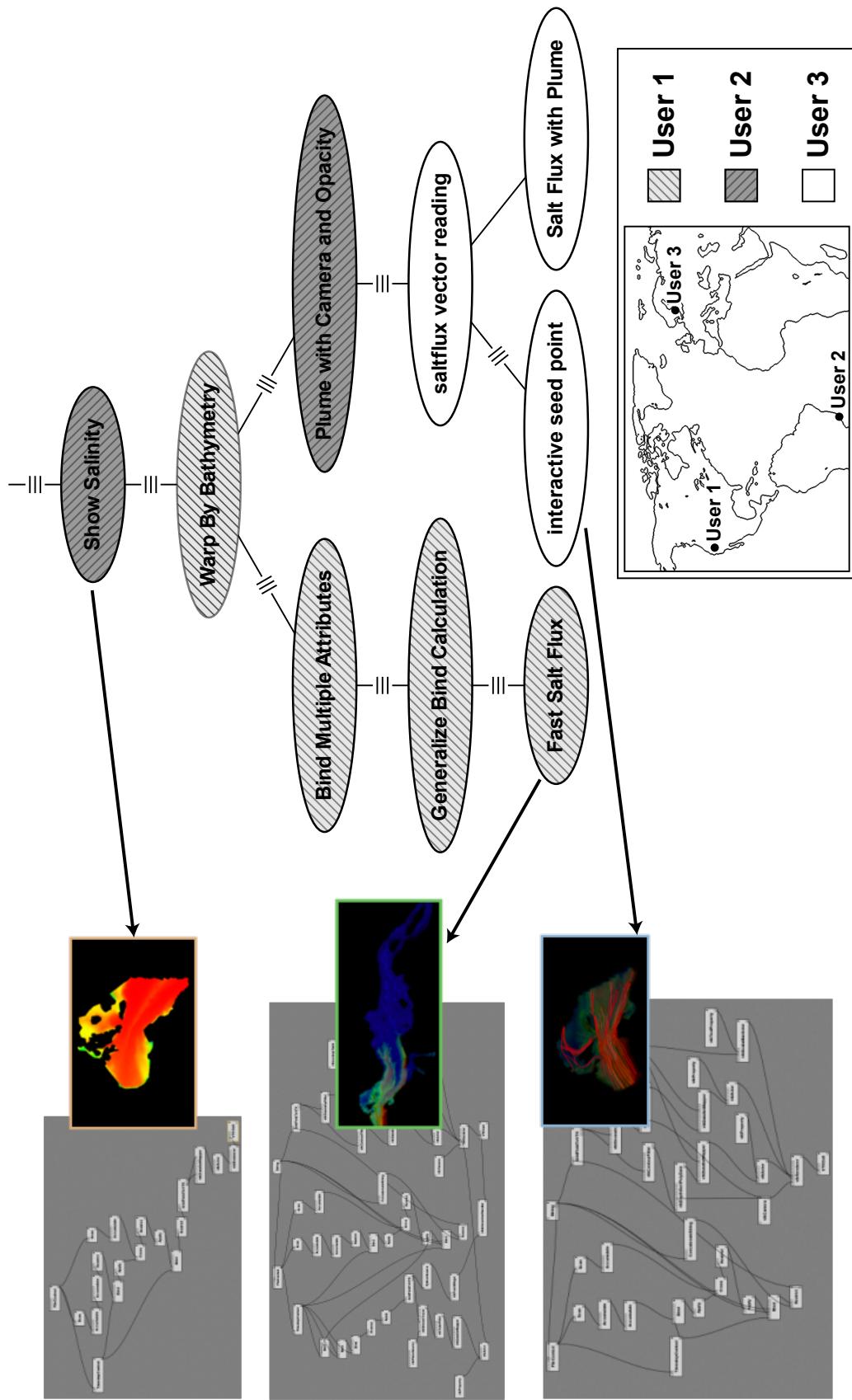
Figure 3.6: A version tree containing a series of workflows that derive visualizations of the Columbia River Estuary. The visualizations have been created by collaborating users. Versions created by different users are represented using different colors.

be the largest action id in a set of actions A. Then, for two sets of actions, $A_1 \subseteq A_2$, the set of new actions, ΔA, is

$$\Delta A = \{a \in A_2 \mid N(A_1) < \mathbf{id}(a) \leq N(A_2)\}$$

This means that we can efficiently determine which actions a user requires to update his version tree. If a user has copied all of the actions in the database up to id N_D, then we only need to copy actions a_i with $\mathbf{id}(a_i) > N_D$ from the database. Conversely, if a user has already saved all actions up to N_L to the database, only actions a_i with $\mathbf{id}(a_i) > N_L$ need to be sent to the database. Figure 3.7 shows a simple example of the steps of the algorithm.

3.2.4.1 Relabeling

Determining the set of new actions is easy when one of the two sets being compared is a superset of the other. However, when multiple users are collaborating, we might not be in this situation. Consider the scenario shown in Figure 3.8, where user A and user B made changes at the same time. Both clients will try to simultaneously save their actions to the database before being notified of the other's changes. In each of their local version trees, they both have actions with id 7, but these actions are not the same. Assuming A's request gets to the repository first, her action will be given id 7 while B's action will become id 8. Thus, after pushing out the other's updates, A and B will have the same tree except that the ids of the nodes may differ.

Since an update of the ids in the local version tree might interfere with a user's current work, we choose to maintain a set of local ids that can be mapped to the global repository ids. Specifically, we maintain a bijective map

$$M : \mathbf{id_{global}} \leftrightarrow \mathbf{id_{local}}$$

Let $M_{\mathbf{local}}$ denote the reverse mapping from global to local and $M_{\mathbf{global}}$ denote the forward mapping from local to global. All user operations will be accomplished using the local ids, but whenever we need to save to the centralized repository, we translate everything to the global set of ids. Figure 3.8 shows an example of this relabeling.

3.2.4.2 Beyond Action

As described earlier, an action contains metadata and a set of atomic operations. The metadata and the atomic operations, in turn, contain their own ids and may also include references to other entities. Thus, the relabeling of an action needs to update these references as well. For example, each action stores both its own id (**action.id**) and its parent id (**action.prev_id**). If we update the id of the action referenced by **action.prev_id**, we also need to update the **prev_id** field. The same is true for child objects. Suppose the connection in an `add connection` operation references the two
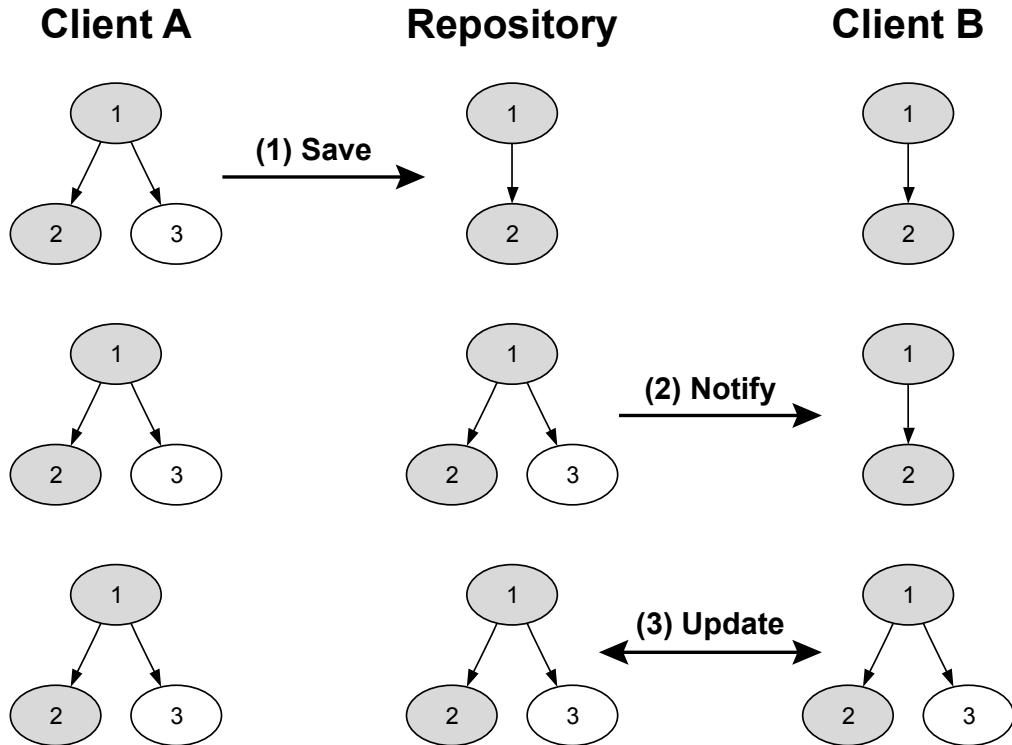
Figure 3.7: The synchronization algorithm. Client A creates a new change (labeled as version 3). This new version is automatically saved to the repository (Step 1). Whenever the repository is updated, it notifies all clients of the new change (Step 2). All clients (including Client B) then incrementally update themselves (Step 3).
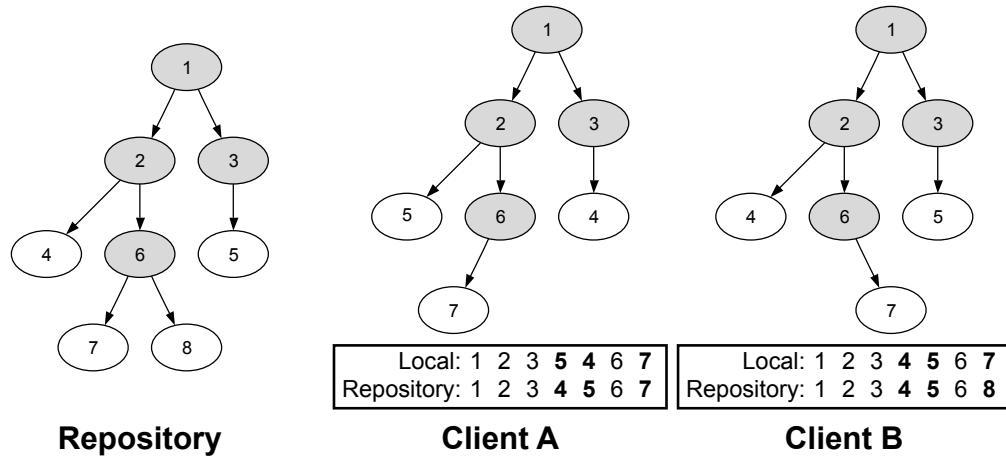
Figure 3.8: Relabeling. Because two users may make updates at the same time or may temporarily lose their connections with the repository, the ids of their nodes may not correspond with the repository's ids. To solve this problem, each client stores the tree according to its own local ids and maintains a map to the repository's global ids.

modules it connects by id. If we remap the id of one or both of those modules in an `add module` operation, we need to update the ids in the `add connection` operation as well. This requires an ordering that respects the properties being updated; we impose an explicit order on modules and connections so that all modules are relabeled before connections to ensure all references are updated.

3.2.4.3 Algorithm Specifics

We combine the method for determining new actions with our relabeling strategy to obtain robust algorithms for incrementally loading from and saving to a database. Algorithm 1 describes the loading algorithm and Algorithm 2 summarizes the saving algorithm. In each algorithm, we use either the database or local version tree to update the other depending on the direction, ensuring that new ids are assigned, existing ids are remapped, and the global-to-local mapping M is updated. Note that all entities are updated in place, copying only the (new) required information from one side to the other.

3.2.5 Implementation

We have implemented the synchronization mechanism on top of the VisTrails system (http://www.vistrails.org). The implementation consists of a client/server architecture shown in Figure 3.7. The server side is a MySQL database that stores version trees. Users can create synchronization sessions through the user interface (see below). The standard VisTrails database schema has been extended to store information about synchronized sessions. This information includes the ids of synchronized version trees, user ids, IP addresses, and port numbers. A database trigger uses this information to notify clients when relevant updates are available. The notification is done by an external MySQL function that uses a socket to connect to the client. The message to the client includes the version tree id number so that the client can request the updates for that version tree. Note that messages about changes to a given version tree are sent to all users using that version tree, except to the user whose changes activated the trigger.

The client side application is a modified version of VisTrails; the modifications include code for performing incremental updates and saves against the database and for receiving notification messages from the database. Because the system contains a controller object for each version tree, we use it to monitor these notifications and start update procedures. Because the controller is linked to the GUI, we also need to redraw the version tree whenever synchronization modifies the tree.

To set up synchronization, users need to select (or create a database) to serve as a centralized repository. This database must have the schema as outlined above and the synchronization triggers that send the update notifications. Once the database is in place, users connect to the database and

Algorithm 1: Incremental Load Algorithm

Input: The local version tree V, $\mathbf{id}_V$ (the id function for V), the global-to-local id map M, and the centralized repository D.

Output: None. It updates both V and M in place.

$\text{LOAD}(V, \mathbf{id}_V, M, D)$

(1)   **max_id** $\leftarrow$ Query V for the maximum id

(2) $A \leftarrow$ Query D for all actions with id $>$ **max_id**

(3)   **foreach** a **in** A:

(4)   Create a', a local copy of a

(5) $a'.\mathbf{id} \leftarrow \mathbf{id}_V(a)$

(6) $a'.\mathbf{prev\_id} \leftarrow M_{\mathbf{local}}(a.\mathbf{prev\_id})$

(7)   Add pair $(a.\mathbf{id}, a'.\mathbf{id})$ to M

(8)   Add a' to V

Algorithm 2: Incremental Save Algorithm

Input: The local version tree V, $\mathbf{id}_D$ (the id function for D), the global-to-local id map M, and the centralized repository D.

Output: None. It updates both V and M in place.

$\text{STORE}(V, \mathbf{id}_D, M, D)$

(1)   **max_id** $\leftarrow$ Query D for the maximum id

(2) $A \leftarrow$ Query V for all actions with id $>$ **max_id**

(3)   **foreach** a **in** A:

(4)   Create a', a global copy of a

(5) $a'.\mathbf{id} \leftarrow \mathbf{id}_D(a)$

(6) $a'.\mathbf{prev\_id} \leftarrow M_{\mathbf{global}}(a.\mathbf{prev\_id})$

(7)   Add pair $(a'.\mathbf{id}, a.\mathbf{id})$ to M

(8)   Add a' to D

select the version trees they want to share. After that, the synchronization (sync) mode can be enabled with the push of a button. From that point on, the version tree will be kept in sync with the central repository and the other users. To help distinguish between versions, those created by other users are shown in blue while a user's own versions are highlighted in orange.

3.2.6   Issues

3.2.6.1   Mutable objects

The monotonicity of the version tree is required for the synchronization process. Change actions and operations are immutable: they are never modified after they are stored in the repository. Thus, the system only needs to check for *new* objects in order to perform synchronization. There are, however, *mutable* objects associated with actions for which this optimization cannot be applied. For example, VisTrails has *version tags* and *version annotations* associated with workflows that

can be modified, and these modifications are not saved as actions. Version tags assign text labels to workflow versions while version annotations store general notes about the version. Because changes to these objects are nonmonotonic (and destructive), *all* objects must be saved and loaded during each incremental load/save. Locally, we can keep a flag that indicates whether or not the entity changed so that we only need to save it when it does, but the same cannot be done for the global repository. Nonetheless, since the volume of mutable data is small, we copy all instances during an incremental load.

3.2.6.2   Integrating changes

One nice feature of our synchronization framework is that it does not require the user to integrate another user's changes. However, consider the situation where two users (A and B) are working on a similar problem, and they have attacked different pieces of it from a common starting point. Each has seen that the other has made changes, but they wanted to finish their own piece. Later, when they decide to integrate these changes, user A can switch to B's version and make the changes applied in her own version. A more efficient alternative would be for user A to use the *analogies* mechanism [136] implemented in VisTrails to automatically apply the changes from one branch to another.

3.2.6.3   Local parameters

Workflows may not always have the same meaning to all users, and they may disagree about certain parameter settings or methods used. For example, an input filename parameter may differ between two users because the users store the file in different disk locations. Currently, the only way to deal with such local parameter settings is to create a different version for each set of parameters. This means that a change in one user workflow will not propagate to the other version, which is not desirable. A solution to this problem could be to separate the shared workflow from the local settings creating a division of the workflow in some way.

3.2.6.4   Data sharing

The ability to share data is an important part of collaboration. For workflows, you may want to share output data as well as input and intermediate results. This can be done with a data pool which maintains up-to-date data items created by the users. This would make it possible for users not only to see each other's results, but also use the data as inputs to other workflows. The COVISA project[162] implements this kind of data sharing. Users can exchange data and directly use them in their pipelines. Another system that implements the idea of a data pool is the *Data Playground*[60]. The Data Playground provides a workflow editor that is highly data centric, letting users view and

import data while they compose workflows that in turn create new data items. This gives the users control over their data while they experiment with different data manipulation operations. The prototype only works for one user but it shows how a data centric view can be used in collaborative workflow design.

3.2.6.5 Module packages

A requirement for users to be able to share workflow specifications is that they both use the same repository of module packages. Module packages contain sets of modules that perform similar functions, much like web services. If one collaborator is missing a module, a workflow containing that module cannot be executed. For collaborations that require many different packages and libraries, an effective mechanism is needed for sharing, for example, through the use of public repositories or automatic methods for users to import module packages from other users as they are required. The packages are often platform specific and versioned, so finding the right package is not trivial. This requires packages to use a good version scheme, with possibly backward compatible packages. There also need to exist different versions for different platforms so that the users platform can be identified and the correct package used. Another way to handle module sharing is to use shared computing infrastructure, such as the TeraGrid (http://www.teragrid.org), which can provide a comprehensive set of packages.

3.2.7 Collaborative Design in Multidisciplinary Research

In today's scientific community, it is rarely the case that novel scientific discoveries can be made by a single person. Unfortunately, in many instances of close collaboration, the various domain experts are unable to work in the same location. These types of relationships benefit greatly from the ability to concurrently modify a given workflow description.

An example of the advantages gained from collaboratively designed workflows can be seen in collaborations between the authors at the University of Utah and researchers at the Center for Coastal Margin Observation and Prediction (CMOP).[1] CMOP scientists, located in Oregon and Washington, often spend a significant amount of time describing the various processing and analysis methods they employ to understand their data. While in many cases e-mail is satisfactory for sharing knowledge with collaborators, in some situations, a more immersive collaborative workspace is required.

When a task relating to a specific researcher's area of expertise is being considered, it is often necessary to synchronize processing workflows to arrive at a desired result. By allowing scien-

---

[1]http://www.stccmop.org

tists at the CMOP centers in Oregon to work synchronously with researchers at the University of Utah, the critical task of communication is enriched. Instead of relying on e-mail and telephone conversations to ask important, and often time consuming, questions, scientists can explore *and fix* each others' processing and parameterization errors in realtime. This behavior is exemplified in Figure 3.9. This degree of collaborative design reduces the number and severity of communication based misunderstandings as well as increases the level of productivity of everyone involved in the project.

There are existing mechanisms that can be used for collaborative design of workflows. One of the most general and common methods of realtime collaboration is through remote desktops like VNC [122]. By using this in the design of a workflow, users can see each others' operations like dragging modules around and creating connections. But for more efficient modes of interaction, both users need to be in control simultaneously, and be able to choose whether to take notice of other users' activities. In addition, provenance information would be lost, since it would not be possible to distinguish changes performed by different users.

A related area is that of collaborative visualization such as the COVISA project[162] and No-CoV [158]. COVISA enables several modes of collaboration like sharing data, sharing control of parameters and instructor driven collaboration where one user is in control of another user's pipeline. NoCoV enables users to collaboratively edit a pipeline consisting of instances of *Notification Web Services*. Both of these systems enable collaboration in the creation of the visualization pipeline but they do not support the exchange or existence of different versions of the pipeline.

The use of realtime collaboration has been explored in other areas. *Cobrowsing* [48] enables multiple people to browse the by sharing a Web browser view and following links together. Similar to VNC, cobrowsing is useful when a user wants to guide another through a browsing session. However, unlike VNC where the whole desktop is shared, in cobrowsing users only share a browser view. Cobrowsing can thus be more efficient, since only clicks within a browser view need to be propagated to the users.

A more indirect way of sharing workflows is through public repositories, like myExperiment [61] and Yahoo! Pipes [163], that have become available recently. These repositories foster the reuse of knowledge. They provide search interfaces that allow the users to locate workflows that solve a particular task, and then integrate these workflows into their own. The synchronization infrastructure we propose could potentially be a useful feature offered by these sites.

3.3   EEG Analysis and Visualization

Workflows created in VisTrails provide more than enough functionality for a robust suite of EEG analysis tools. Additionally, VisTrails allows large and complex workflows to enable interaction
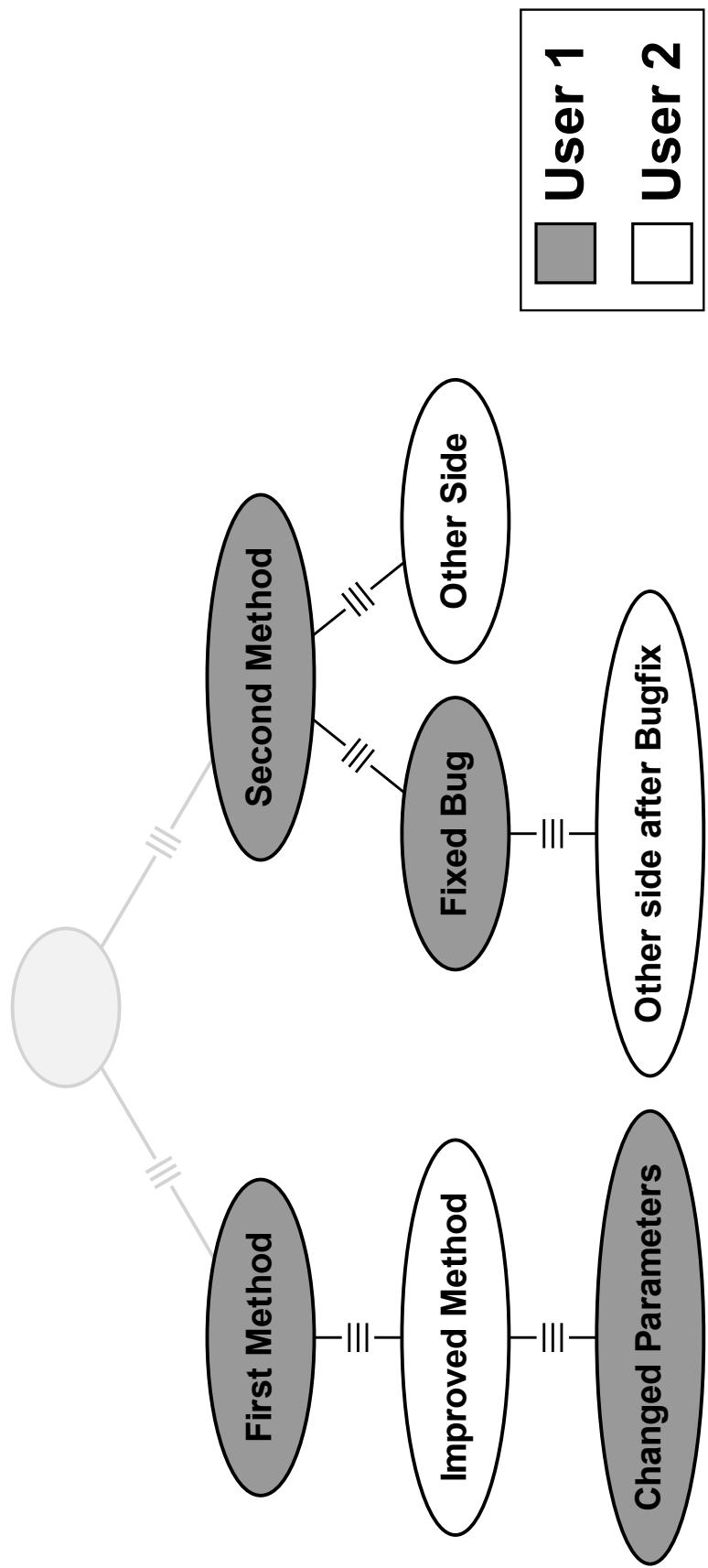
Figure 3.9: An example of collaborative design. Here, two persons have built on each others' workflow specifications, leading to incrementally better results.

and comparative visualization. As in other systems, this flexibility often introduces unnecessary difficulties for users of the tool. To address this extra layer of complexity, a package of VisTrails modules has been created to simplify the construction of analysis and visualization workflows.

3.3.1 Signal Processing in VisTrails

VisTrails' ability to interact with multiple libraries and programming languages through Python is an important advantage to the use of the system. The NumPy (`http://numpy.scipy.org`) and SciPy (`http://www.scipy.org`) Python packages give Python the support to handle not just arrays, but matrices through a collection of robust processing and manipulation routines. Many of the routines provided by SciPy, particularly those in the signal processing toolkit, are computationally expensive. These more costly algorithms are implemented outside of the Python environment in high performance languages such as C and FORTRAN. The NumSciPy VisTrails package wraps these Python packages to expose their functionality within VisTrails.

Several signal processing techniques useful for the analysis of EEG data are implemented in high performance environments and exposed for use in Python. Optimized Fast Fourier Transforms (FFTs) and Short Time Fourier Transforms (STFTs) are provided by the SciPy package. Additionally, numerous other packages are available to Python not only by SciPy, but also through the wrapping of external libraries facilitated by *ctypes* package (`http://python.net/crew/theller/ctypes`) as discussed by Anderson et al. [8].

Of particular interest to EEG processing is the Stockwell Transform [147]. This time frequency transform is used in spectral analysis as it is similar to the continuous wavelet transform yet maps directly to the complex frequency domain. This functionality is made available to VisTrails by leveraging ctypes with NumPy array structures. Combining C-style arrays with Python in this way enables easy integration of new functionality with the existing NumPy and SciPy module packages in VisTrails.

Figure 3.10 illustrates some of the fundamental differences between the STFT, (Continuous) Wavelet Transform and the Stockwell Transform. While the STFT offers a mapping to the complex Fourier Domain, the resolution of the transform is determined solely by the support of the windowing function applied. This dependence results in a uniform tiling of the time frequency space. Unline the STFT, the Continuous Wavelet Transform provides an adaptive tiling of the time frequency domain, resulting in more optimal resolutions in both time and frequency. However, this transform is a mapping not to the complex Fourier domain, but to a time-scale space in which *pseduofrequency* is determined by the scale of the wavelet basis function, and the notion of phase is not well described. The Stockwell Transform, on the other hand, takes advantage of the direct

(a) (b) (c)
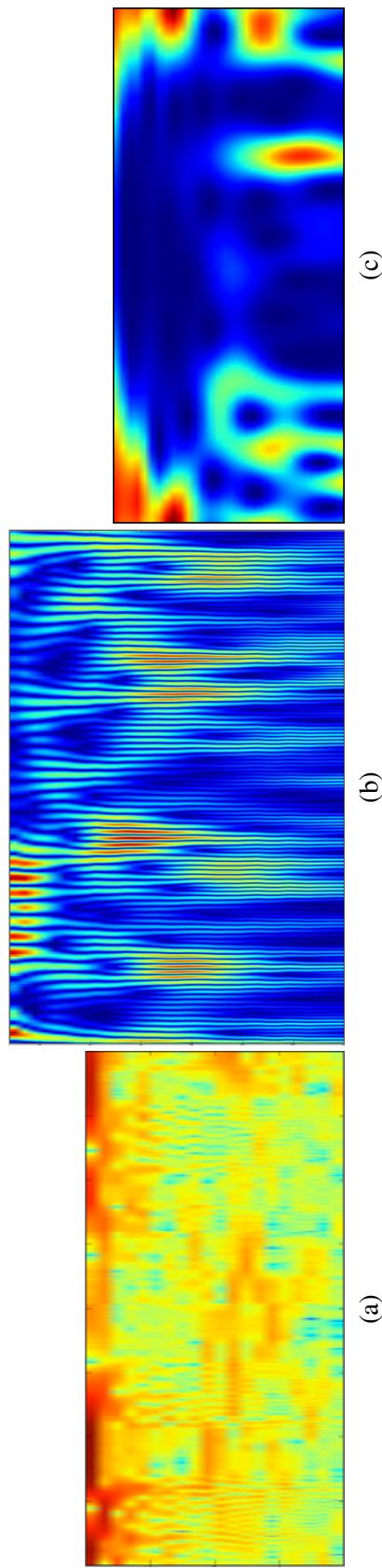
Figure 3.10: The resolution of the Short Time Fourier Transform (STFT) depends on the support of the windowing function being used, resulting in a uniform tiling of the domain (a). However, the Continuous Wavelet Transform uses an adaptive resolution scheme to more optimally represent data, but does not represent data in a Fourier like domain (b). The Stockwell Transform maintains a mapping to the Fourier domain while using an adaptive resolution scheme similar to the Wavelet Transform (c).

mapping of a time series to the complex Fourier domain as well as the adaptive resolution properties of the Continuous Wavelet Transform [147].

3.3.2   Visualization in VisTrails

Visualization in VisTrails is accomplished through the use of specialized visualization libraries such as the Visualization Toolkit (VTK) [138]. The dataflow paradigm used in VTK is similar to that of VisTrails and enables a straightforward translation into the module based design of VisTrails packages. Unfortunately, the size and complexity of VTK represents a substantial barrier for novice users. To address this issue, the VisTrails Brainiac package utilizes subworkflows to abstract the details of VTK rendering methods into simple, easy to use single Modules.

In addition to VTK based visualization, other plotting and rendering libraries are easily incorporated into VisTrails. Matplotlib [74] provides plotting functionality for use with native Python datatypes as well as NumPy arrays. Matplotlib's scripted nature also exposes the full gamut of plot parametrizations to the user. However, the flexibility afforded by Matplotlib is often cumbersome when it is not needed. To address this problem, Modules in the NumSciPy VisTrails package have been created to wrap the most commonly used plotting functionality within a single Module.

The VisTrails spreadsheet is also amenable to comparative visualization and analysis [9]. By giving users access to a robust spreadsheet on which results are displayed, side-by-side comparisons become possible. The direct comparison of results in a spreadsheet interface is particularly useful when assessing the strengths and weaknesses of different processing techniques and when visualizing multiple data products as seen in Figures 3.11 and 3.12.

3.4   VisTrails as an EEG Processing Environment

The field of Neuroscience often uses both multimodal data as well as computationally complex algorithms to analyze data collected from participants in a study. Here, we performed a study in which Magnetic Resonance Imaging (MRI) is combined with Electroencephalography (EEG) to examine working memory performance. The MRI provides a three-dimensional depiction of the structure of the brain and presents a natural spatial organization for the EEG sensors. EEG data are collected from 64 sensors placed on the scalp. These sensors measure the voltages at the scalp generated by brain activity.

The use of MRI data in conjunction with the known EEG sensor locations highlights spatial relationships between the sensors and the brain activity they measure. During more complex analysis, MRI data is often used to determine participant specific finite element meshes for use in solving the inverse problem, assigning scalar values to the cortical surface based on data collected at sensor locations more robustly than interpolation schemes. Solving the inverse problem in this way
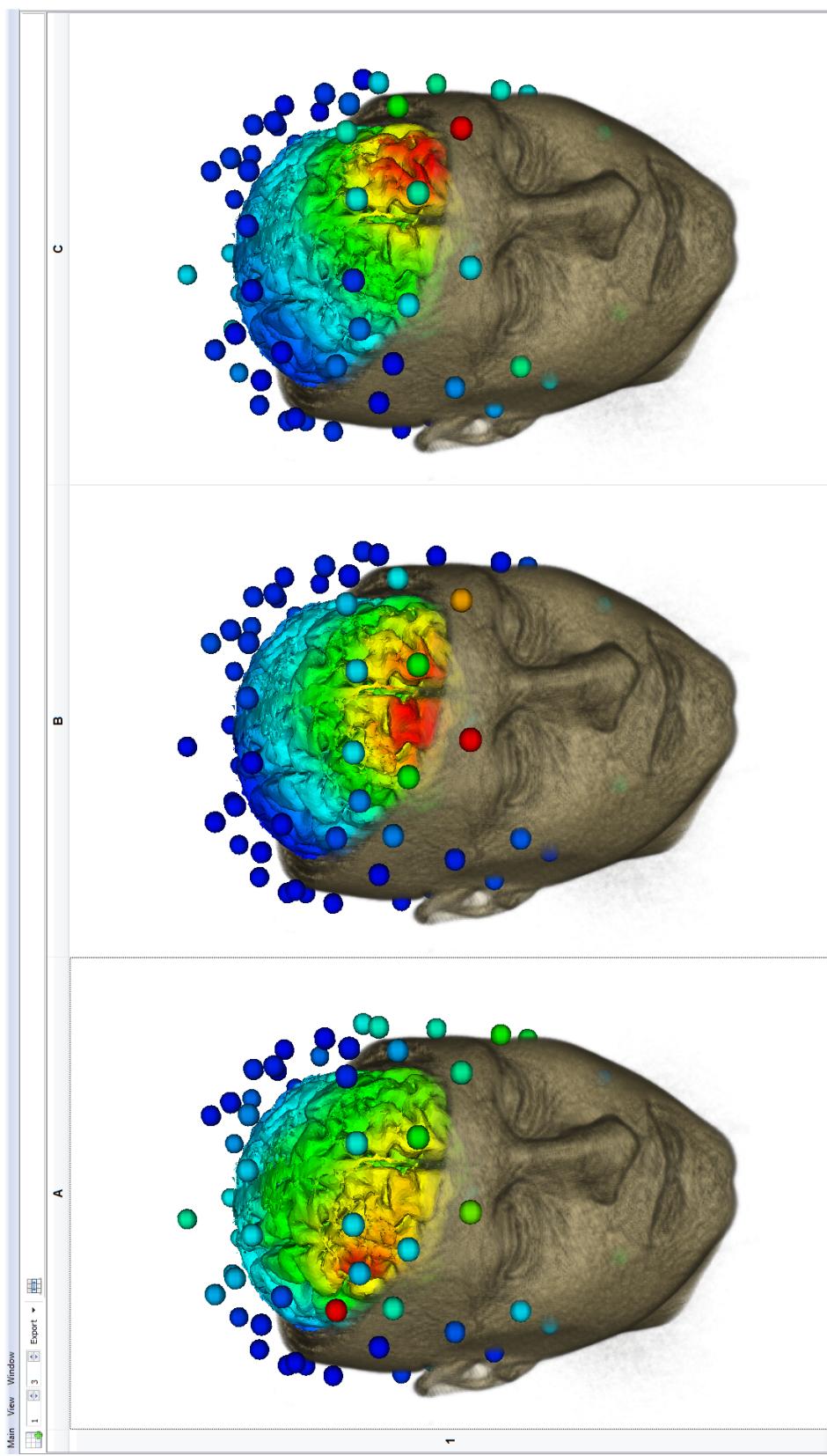
Figure 3.11: Comparative visualization in VisTrails is possible through the Spreadsheet interface. Here, EEG data are processed to extract power values in the alpha band of frequencies and visualized at both the sensor locations as well as the cortical surface via interpolation. These visualizations represent, from left to right, 500 ms, 800 ms, and 1300 ms after trial onset.
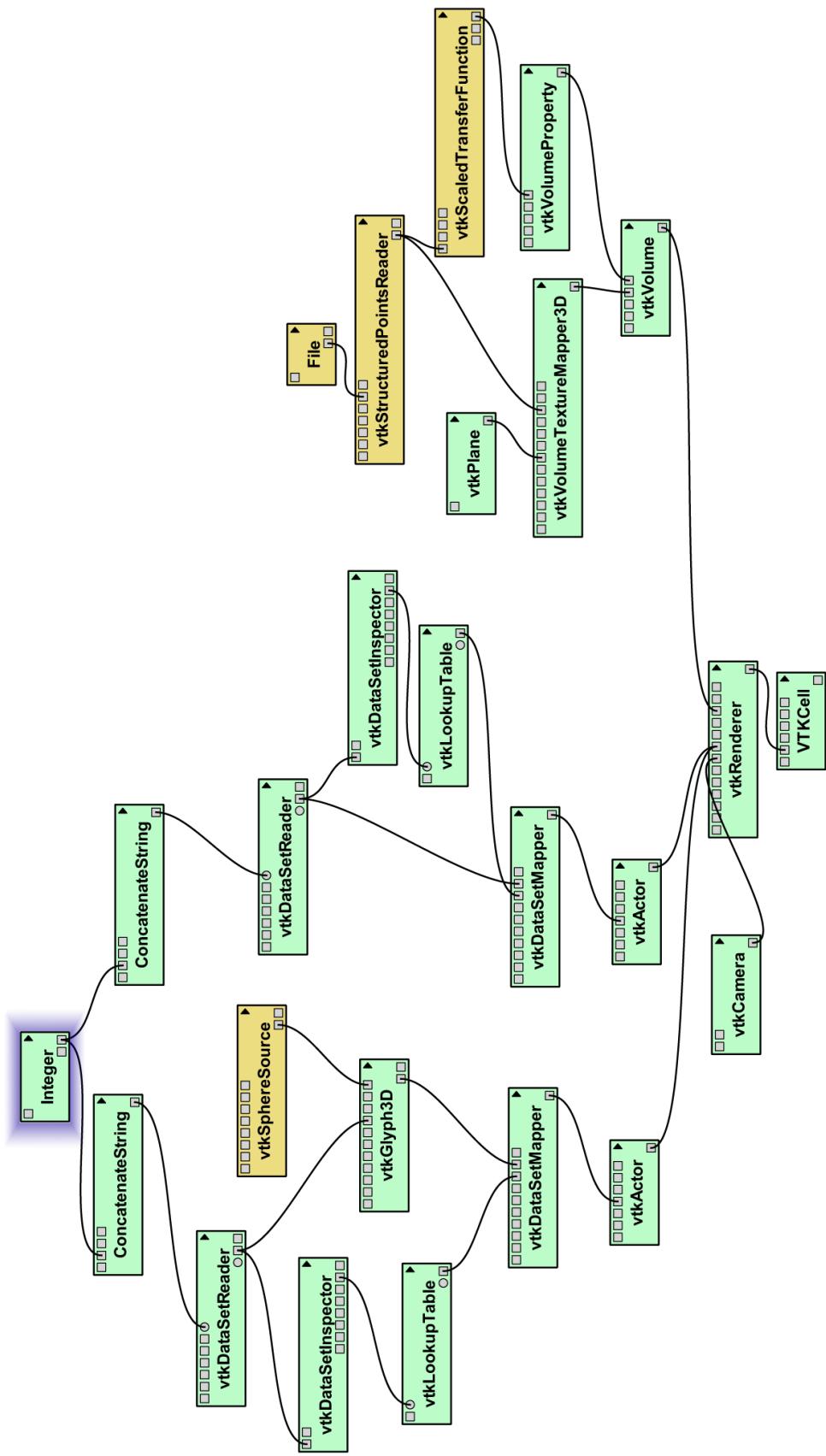
Figure 3.12: The three timesteps visualized in Figure 3.11 are generated by changing the appropriate parameter (highlighted in blue) in the underlying VisTrails Workflow.

is useful for tasks such as source localization for epileptogenic regions of the brain. Fortunately, for determining spatial relationships between active brain regions and EEG sensors, radial basis function interpolation has proven to be a good approximation [108]. Regardless of the method by which scalars are mapped from discrete point locations to a surface representation, data collected from different acquisition methods must be fused into a coherent representation.

For the type of data fusion necessary for advanced visualization of EEG data to be performed, the MRI data must first be registered with the EEG sensor locations, as seen in Figure 3.13. In this case, sensor locations were provided by the hardware vendor in normalized Talairach Space [149]. To properly align sensor locations with anatomical fiducials, the MRI volume was transformed in an offline process to ensure that both sensors and participant specific MRI data were described in the same coordinate system. Registration of the EEG sensors with the MRI data is important for several aspects of the visualization process: the visualization of sensor locations provides additional context when interpreting the rendering, and registered sensor locations are a necessary input when mapping data from individual sensors to the cortical surface.

The transformation of MRI data into Talairach Space is also advantageous for segmentation [49]. In VisTrails, the functionality of the Insight Toolkit (ITK) is leveraged to enable on the fly segmentation of MRI data. Figure 3.14 shows an example workflow used to segment and render the cortical surface extracted from an MRI volume. The figure highlights the flexible nature of VisTrails and its ability to use multiple, independent libraries within a single workflow. Here, ITK performs 3D segmentation to extract the cortical surface while the Visualization Toolkit (VTK) volume renders the result to validate its performance.

Visualizing EEG data while respecting the spatial positions of the electrodes is important when examining activity across the cortical surface. Methods such as topographic maps [108] use simple, two-dimensional representations of a head to provide spatial cues for the visualization, as seen in Figure 3.15. However, VisTrails' abilities to use multiple, external libraries in concert make more advanced visualizations possible.

The use of MRI data with the known EEG sensor locations highlights spatial relationships between the sensors and the brain activity they measure. During more complex analysis, MRI data are often used to determine participant specific finite element meshes for use in solving the inverse problem, assigning scalar values to the cortical surface based on data collected at sensor locations more robustly than interpolation schemes. Solving the inverse problem in this way is useful for tasks such as source localization for epileptogenic regions of the brain. Fortunately, for determining spatial relationships between active brain regions and EEG sensors, Radial Basis Function (RBF) interpolation has proven to be a good approximation [108]. Figure 3.15 compares visualizations
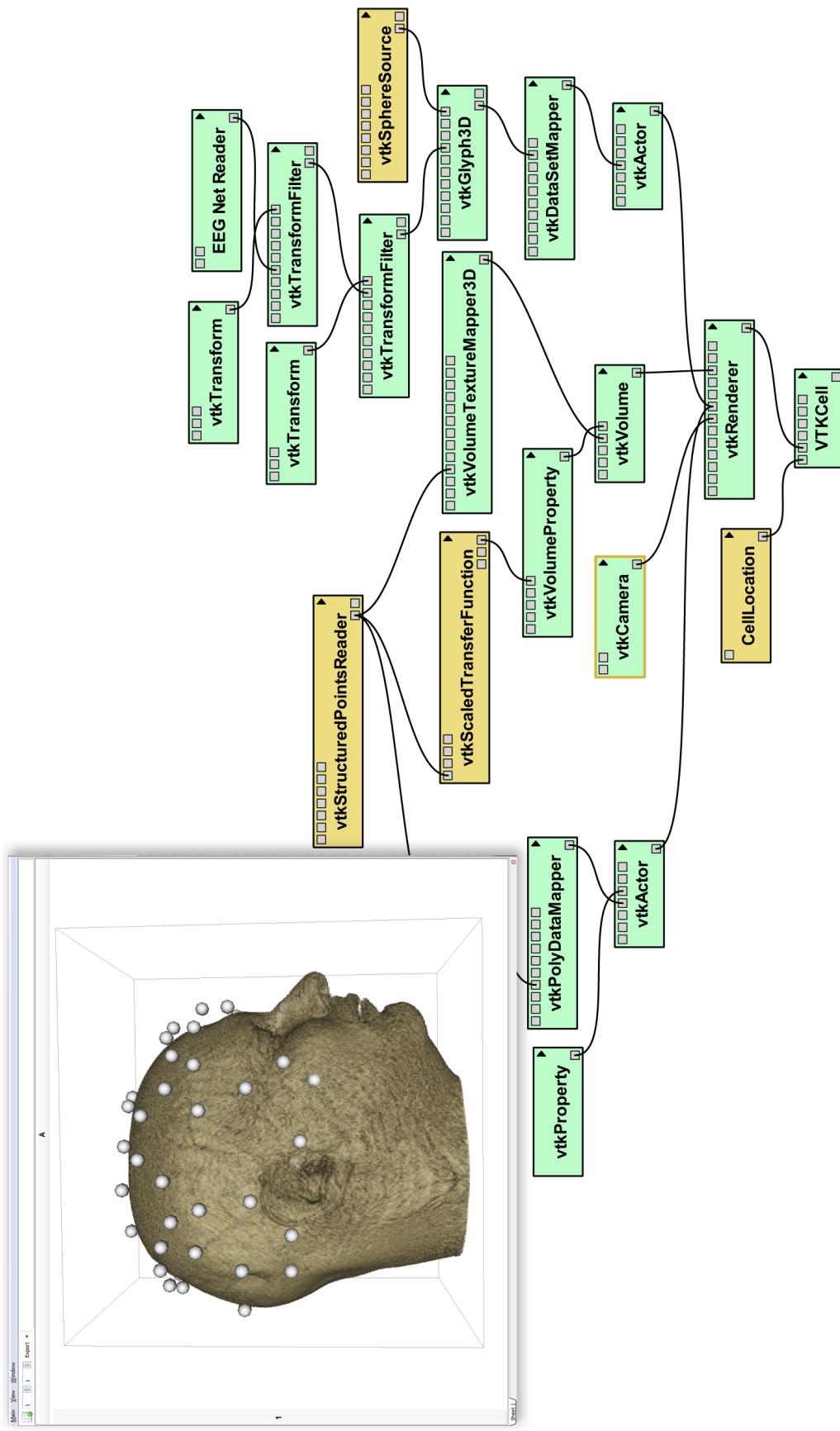
Figure 3.13: Visualizing EEG sensors along with participant specific MRI data (inset) provides additional spatial cues to better understand EEG results. The VisTrails workflow used to generate this rendering takes advantage of the offline registration of sensor location to MRI data and uses only affine transformations to approximate absolute sensor positions. It is worthwhile to note that sensor localization in this manner is an approximation; it is possible to record precise sensor locations at data acquisition time by using additional hardware.
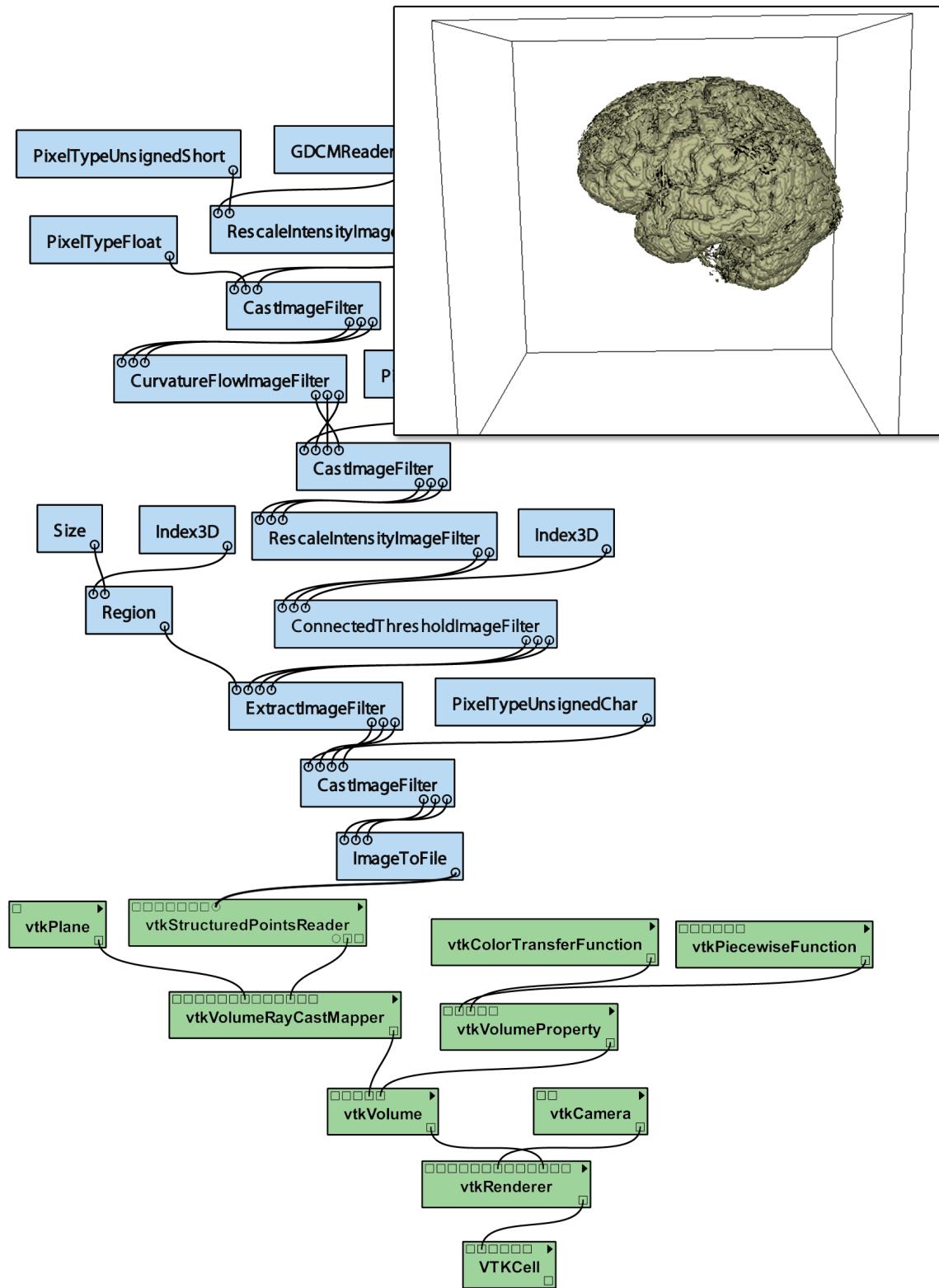
Figure 3.14: The ability of VisTrails to draw upon the features and capabilities of many libraries is useful when segmenting MRI data. Here, the cortical surface is segmented from an MRI volume using the Insight Toolkit (ITK) and then rendered (inset) using the Visualization Toolkit (VTK). Modules in blue represent ITK functionality, while those in green are from the VTK package.
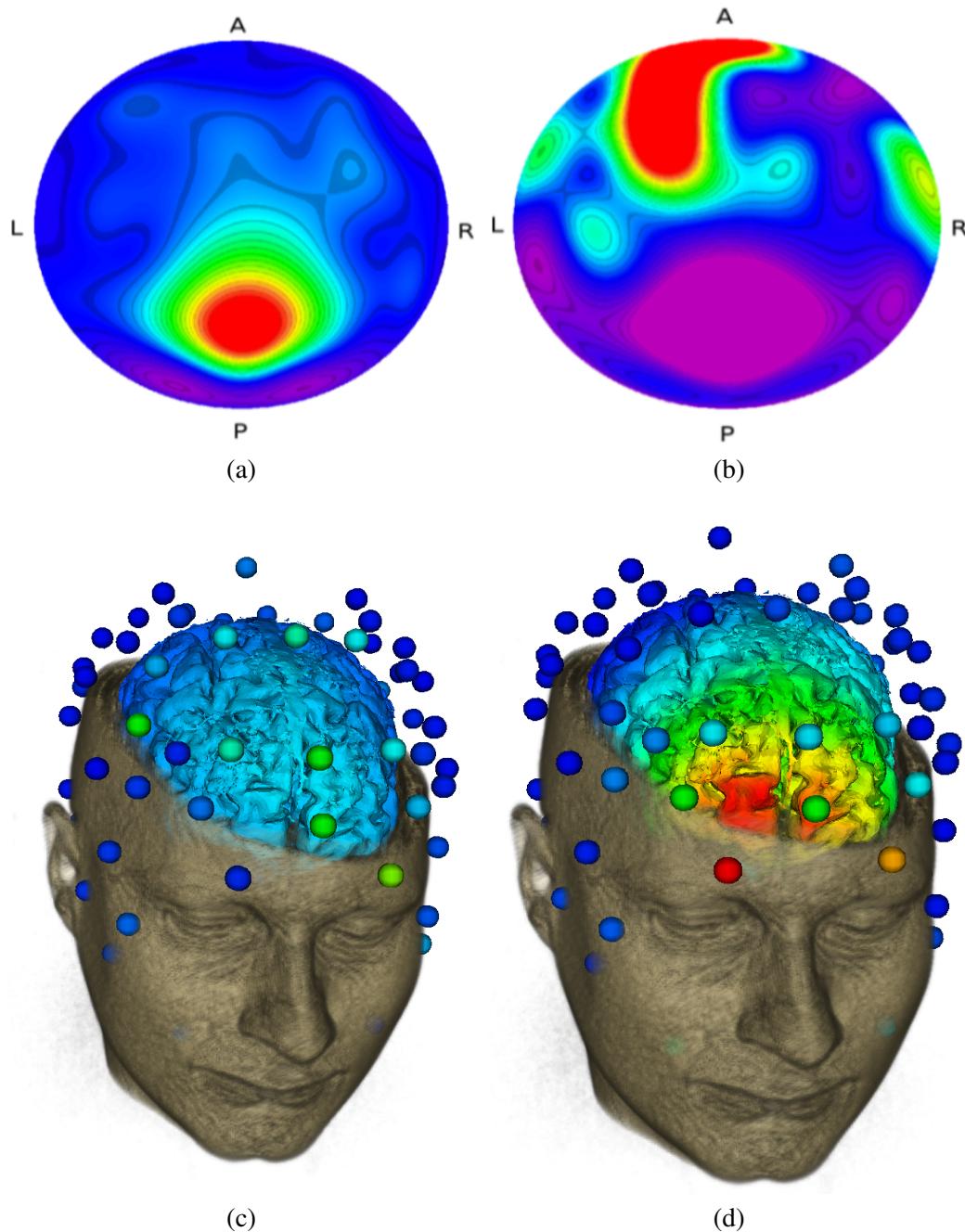
Figure 3.15: Topographic maps representing a resting period (a) and stimulation period (b) are compared with three-dimensional renderings mapping scalars to the cortical surface extracted from MRI data (c,d). Spatial context is better preserved in the three-dimensional rendering and enhanced by interactive capabilities.

using topographic maps with those formed by using participant specific data directly, both of which use RBF interpolation to map scalar data from the sensor locations to the domain representing the head or cortex.

Although both topographic maps and 3D renderings using RBF interpolation provide good spatial cues for understanding brain activity measured by EEG, it is important to inspect the raw data from each sensor as well as their respective time frequency transforms. To appropriately explore data in this way, VisTrails uses VTK's interaction abilities combined with the computational abilities of Numpy and Scipy. Figure 3.16 shows how the various interactive elements of VTK may be used. Users may interact with a clipping plane to expose more or less of the cortical surface by removing portions of the MRI. Additionally, each EEG sensor is selectable allowing users to inspect the raw data streams produced at each electrode. In this way, a more complete picture of brain activity is formed, allowing more robust interpretations by scientists.

3.5 Discussion

Tools allowing rapid exploration of large and multimodal datasets are more important than ever in scientific research. Interpreted languages, like Python, provide a solid foundation for the development of powerful, yet flexible data analysis and visualization tools. However, flexibility of analysis and visualization must be combined to enhance the exploration process. The VisTrails system is a visual programming paradigm in which computational elements are represented by drag-and-drop modules that are connected together to form programs. The drag-and-drop system makes replacing functionally equivalent computations, such as replacing an STFT with a Stockwell Transform, easy to do.

Providing a tool that supports flexible visualization and analysis allows scientists to draw more insightful conclusions. Additionally, the ability to change analysis techniques enables important insights to be gained more quickly. Using visual programming paradigms, like VisTrails, makes changing analysis techniques easier for nonprogrammers, facilitating the use of the tool for insight generation.

The VisTrails infrastructure can be integrated with any workflow system that captures workflow evolution provenance. Our implementation of the synchronization mechanism on top of the Vis-Trails system shows that workflow systems can be a powerful tool for realtime collaboration. Users can collaborate efficiently and effectively, exploring different branches and taking advantage of each others' progress. Together with techniques for data sharing and remote execution, this enables efficient creation of complex workflows.

By leveraging the concise representation of workflows provided by the change based provenance model, synchronization is efficient: only incremental changes need to be propagated to collaborating
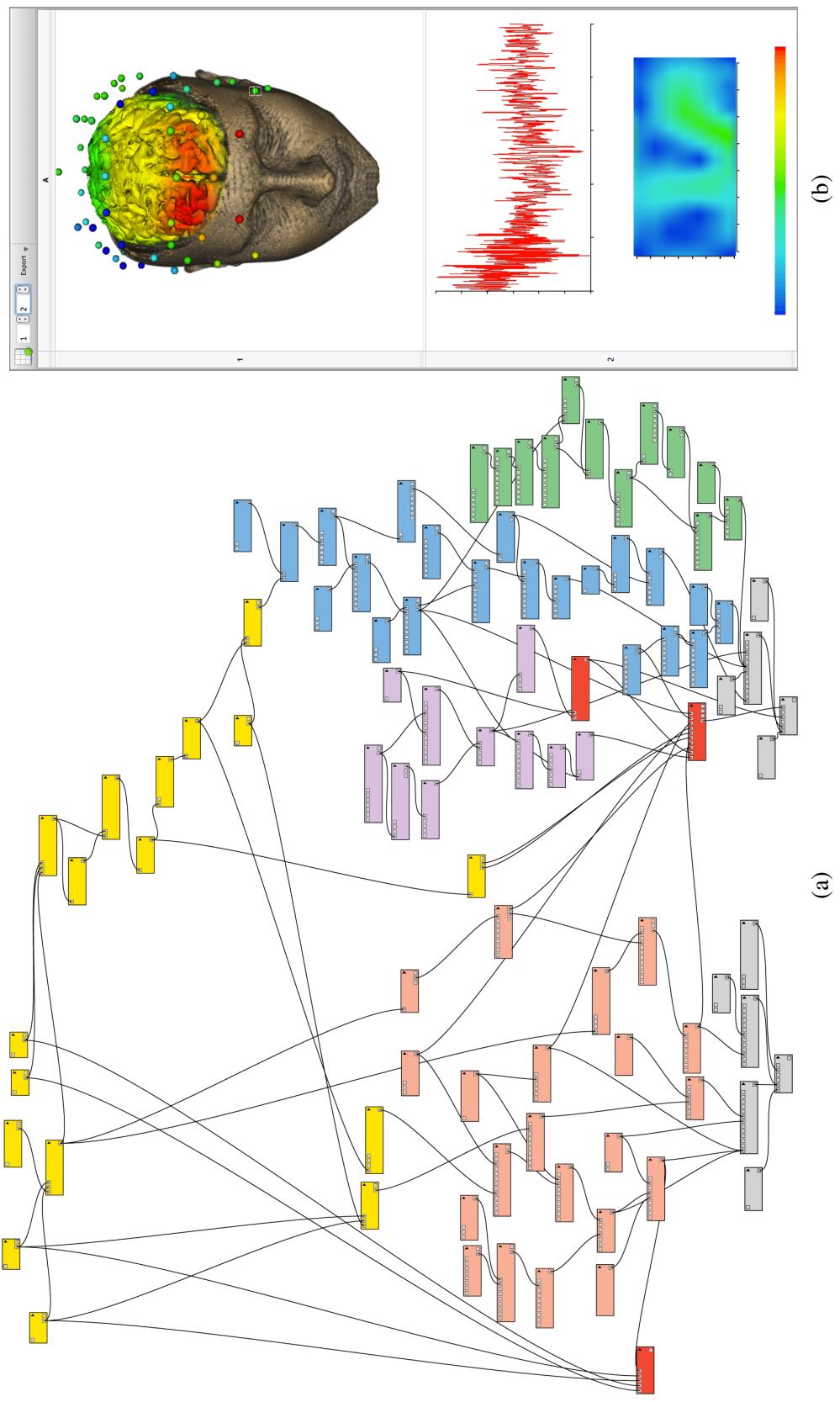
Figure 3.16: VisTrails workflows often contain groups of modules from different libraries. The workflow depicted here (a) loads and processes EEG data with the NumSciPy VisTrails Package (yellow), before using the results to color spheres representing EEG electrodes (blue), the preprocessed cortical surface (green), and raw sensor displays (orange). These data are then combined with a volume rendering of the MRI data (purple) and displayed side-by-side with the raw EEG data and their time frequency decomposition (grey). The results of this workflow are presented in two linked views in the VisTrails Spreadsheet (b).

users. However, further experiments are needed to assess the scalability of the current implementation.

We believe that our provenance based synchronization mechanism can be applied to applications other than workflows. Combined with techniques to visualize provenance information, this mechanism can serve as a powerful platform for collaborative design in general. Users can share their work effectively while inspecting each others' contributions. The application of our synchronization infrastructure in other areas of computational design is a direction we plan to pursue in future work.

Facilitating collaboration between VisTrails experts and domain scientists reduces the amount of time to arrive at insightful understanding. Realtime collaboration allows remote collaborators an additional layer of infrastructure to be leveraged during the formation of processing, analysis and visualization workflows. This type of collaboration helps to eliminate more of the barriers associated with long distance collaboration.

In addition to making changes to analysis techniques easier, Python has also proven to be excellent at combining related, yet disparate data into a single, useful representation. This data fusion is exemplified in Figure 3.16 which has successfully combined structural volumetric data with processed time series data. Analysis taking spatial data into account is common in neuroscience; however, most of these techniques respect spatial relationships by scientists selecting groups of sensors located at specific places to focus their analysis efforts. Providing neuroscientists the ability to examine their dataset as a whole allows insightful analysis to be performed more quickly. The efficiency gained by better utilizing visualization as a tool stems from new and unexpected behavior being identified more easily.

CHAPTER 4

MEASURING, MANIPULATING, AND VISUALIZING WORKING MEMORY PERFORMANCE

This chapter is organized as follows: Section 4.1 outlines the functional role of working memory in the human brain, with particular attention paid to its part in neurological disorders. After discussing the means by which working memory may be measured and manipulated in Section 4.2, an experiment designed to validate these methods is outlined in Section 4.3. Section 4.4 then presents the results of the experiment. Section 4.5 then outlines a method by which EEG data in their entirety may be queried and visualized followed by a discussion of the experiment and visualization technique in Section 5.5.

4.1 Working Memory Dynamics

To be able to properly assess the spectral evolution of EEG associated with working memory, each trial in an experiment is processed with respect to its own intertrial rest period. The individual alpha (8–12 Hz) and theta (3–7 Hz) frequencies are determined for both the trial and rest period and their amplitudes measured [78]. By comparing these values, a shift of both the individual frequencies as well as their amplitudes are revealed. The degree of change in these amplitudes, weighted by the amount of shift in the frequency domain, determines the working memory characteristics for each single trial.

4.1.1 Working Memory in Neurological Disorders

Schizophrenia has been recognized for more than a century as a psychotic disorder in which psychosocial recovery is uncommon, principally due not to psychotic symptoms which are relatively easy to treat, but to disabling impairment of cognitive functions first recognized by Emil Kraepelin in 1893 [87]. Kraepelin described poor outcomes in social functioning, associated with impairments of attention, motivation, problem solving, learning, and memory as the principal features of this disorder. Kreapelin's assessment still holds — less than 15% of patients with schizophrenia recover [27, 99]. There is still no treatment for cognitive dysfunction seen in schizophrenia or other

neuropsychiatric disorders [67, 102, 123]. Recent reports from the NIMH supported MATRICS researchers regarding the development of a consensus for a standardized battery of neuropsychological instruments to identify and measure the principal cognitive deficits of schizophrenia reinforce the importance of cognitive impairment in schizophrenia [64]. With regard to our study, it is also notable that psychomotor processing speed measured as reaction time received a strong endorsement for inclusion in the MATRICS test battery.

On the molecular scale, emerging evidence supports a model of cognitive impairment in schizophrenia that is associated with a complex array of polymorphic alleles. Such variants have been shown to be associated with inheritable impairments of cognitive function [30, 119, 139]. So far, attempts to develop new pharmacologic agents based on these important genetic findings have not been successful. In this regard, our plan to develop a macroscalar (neural circuit) approach to impaired memory using alpha frequency (10Hz) rTMS is supported by evidence showing a predictive relationship between EEG alpha band power and working memory task performance [11, 35, 78, 135]. Evidence to support the relationship between cognitive deficits in patients with schizophrenia and low alpha spectrum power in the Dorso Lateral Pre-Frontal Cortex (DLPFC) of first episode, neuroleptic naïve patients with schizophrenia is also supported by the study of Ford and colleagues [50] who found reduced frontotemporal alpha band coherence in an auditory event related potential study comparing healthy subjects with patients. Finally, the study of Klimesch [79] lends strong support for this work. These investigators compared alpha frequency rTMS, sham and 20Hz rTMS, directed to the frontal, parietal or occipital cortices. Important to our goal of developing a circuit based treatment model for impaired memory, studies show that compared to healthy subjects patients with schizophrenia have lower prefrontal alpha band power and lower peak alpha frequency [50].

4.2   Measuring and Manipulating Working Memory

Frequency and target specific effects have been observed in studies of cognitive domains including working memory: for example, enhancement of picture naming and posterior superior temporal cortex [105]; enhancement of episodic memory and left inferior PFC [83]; analogic reasoning and left DLPFC [20]; mental rotation of 3D objects and mesial frontal and right parietal cortex [79]. Finally, Luber et al. showed that TMS could enhance working memory [94]. They predicted that healthy subjects performance on a delayed match-to-sample task would be enhanced when TMS of 1, 5, or 20 Hz was applied to either left dorsolateral prefrontal or midline parietal cortex during the delay phase of the task. They found 5 Hz, but not 1 Hz or 20 Hz, stimulation to the midline parietal site during delay resulted in a significant decrease in RT without a corresponding decrease in accuracy. In a second experiment, 5 Hz rTMS was applied to the parietal site during the delay or during presentation of the probe. Reaction time enhancement occurred only with stimulation during

the delay phase. The authors concluded TMS could enhance working memory performance under specific conditions of time, target, and stimulation frequency.

$$Var(\Theta) \quad = \quad 1 - \frac{1}{n}\sqrt{(\sum_{i=1}^{n} cos(\Theta_i))^2 + (\sum_{i=1}^{n} sin(\Theta_i))^2} \qquad (4.1)$$

$$\omega \quad = \quad \frac{1}{Var(\Theta)} \qquad\qquad\qquad (4.2)$$

We determine the power spectral density of the pre- and poststimulation preprocessed EEG data by first computing the power spectra of the signal extracted in the time domain. The Fast Fourier Transform was applied first to yield a phasor at each wave number. We then determined the magnitude of the phasors in order to ascertain power at a specific frequency. Phase was extracted from the directionality of the phasor. The initial phase analysis provided a phase histogram that was used to derive circular statistics. Care was taken to calculate such statistics while respecting the periodicity of the phase domain. The expression in equation (4.1) was used to compute phase variance; the expression in equation (4.2) was used to compute the phase synchrony index (PSI). Figure 4.1 illustrates the effect of 10 Hz rTMS on phase synchronization in a representative subject that was actively stimulated. Note the difference of poststimulation PSI for active compared with sham 10Hz rTMS.

Figure 4.2 illustrates the power spectrum of the alpha band frequencies in the same region of the DLPFC before and after 10 Hz rTMS. After stimulation, the same cortical area experienced a nontrivial increase in mean alpha power and an increase in alpha peak frequency. Increases in power and shift of the alpha peak are associated with a shorter reaction time, as reflected in Table 4.1. Finally, our data also show an association between active stimulation and the degree of induced alpha band phase synchrony in the stimulated region. Sham stimulation was was not associated with perturbations of alpha power, frequency, synchrony, or improved performance.

4.2.1 The Sternberg Paradigm

Task stimuli and computation of reaction timess and accuracy were derived using Presentation 9.90 software (Neurobehavioral Systems, San Francisco, CA). A single trial of this task was composed of four phases: stimulus (a 5- or 7-item string of uppercase consonants and vowels was presented at the center of a 19 inch video monitor for 1.5 sec), delay (participants focused on a fixation marker appearing at the center of the computer monitor while maintaining the previous string in memory for 1.5 sec), probe (a single uppercase letter appeared at center screen for 3 sec), and response (participants were instructed to decide as quickly and accurately as possible whether the probe was present in the previous string, and to respond by pressing one of two keypad buttons; if
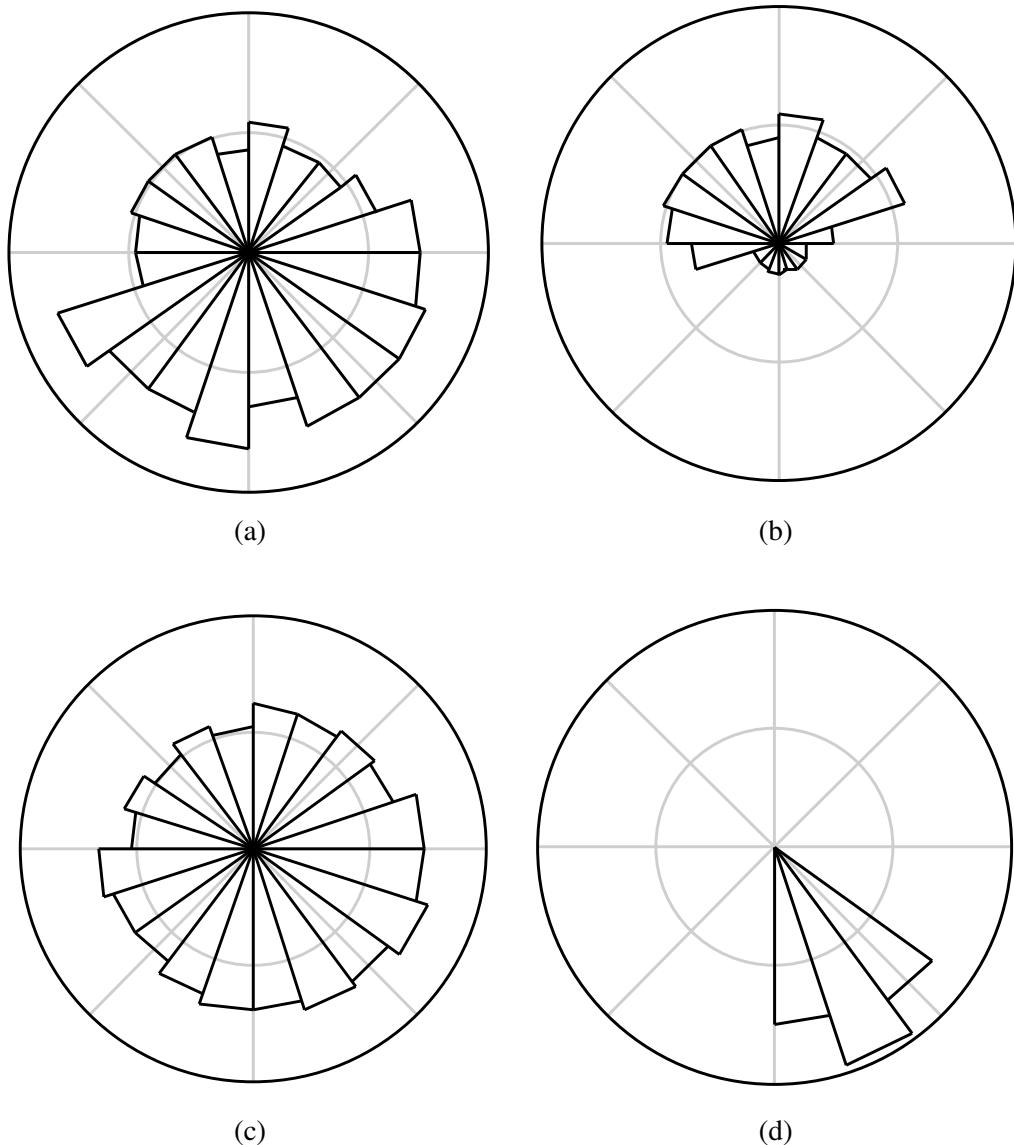
Figure 4.1: Comparison of pre- and poststimulation phase plots support the prediction that 10 Hz rTMS directed to the prefrontal cortex induces alpha frequency synchronization while sham stimulation shows no such effect. The phase histograms of both the sham and active prestimulation groups ((a) and (c), respectively) show a uniform distribution of phase at rest. After stimulation, the sham group shows a placebo effect (b) while the active group exhibits a drastic change in phase synchrony (d).
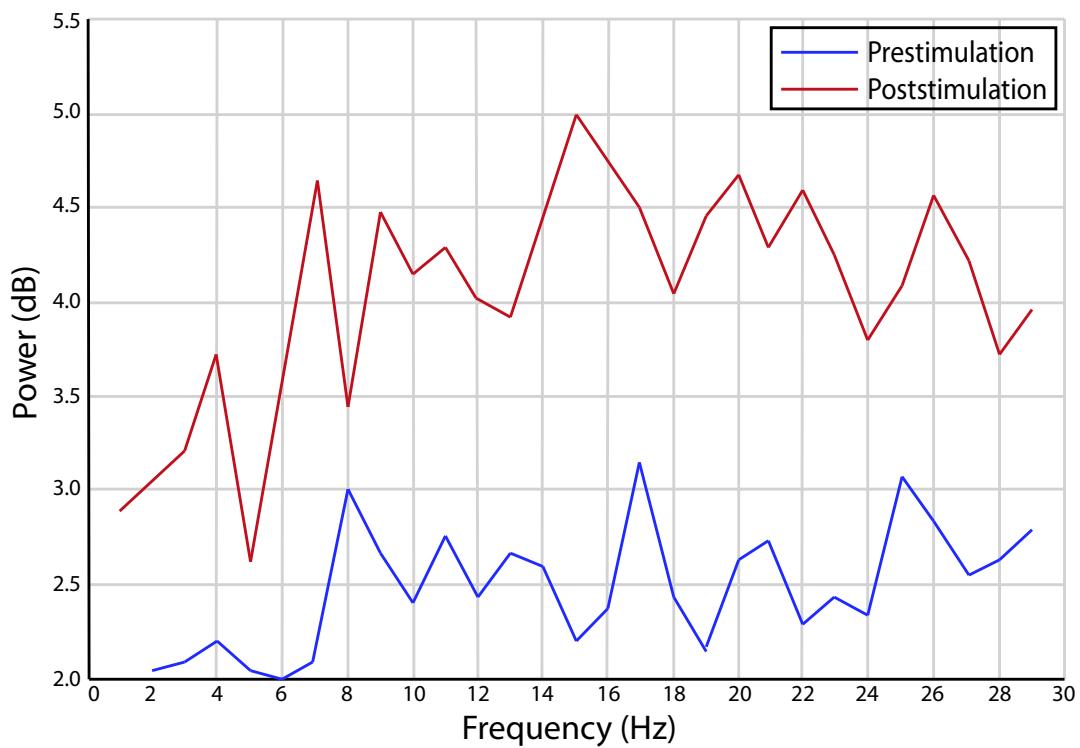
Figure 4.2: Prefrontal cortical broadband power spectral density in a subject in the active 10Hz rTMS group exhibits an apparent increase in broadband power with stimulation. 10Hz rTMS may also perturb alpha mean frequency: note the poststimulation induction of an alpha-1 peak not present prestimulation.

Table 4.1: Sternberg task reaction time pre- and poststimulation with 10 Hz repetitive transcranial magnetic stimulation by group. These data show that working memory is robustly manipulated by application of rTMS.

	δF	F	*Sig*
PreStim RT: *BetweenGroups*	1	0.720	0.407
WithinGroups	18		
PostStim RT: *BetweenGroups*	1	5.700	0.028
WithinGroups	18		
RT Difference: *BetweenGroups*	1	15.262	0.001
WithinGroups	18		
Left PostStim RT: *BetweenGroups*	1	7.378	0.014
WithinGroups	18		
Left RT Difference: *BetweenGroups*	1	17.676	0.001
WithinGroups	18		
Right PostStim RT: *BetweenGroups*	1	3.981	0.061
WithinGroups	18		
Right RT Difference: *BetweenGroups*	1	11.388	0.003
WithinGroups	18		

there was a response within 3 sec, it was recorded, and the next trial began. If there was no response, the trial was marked as an error, and the next trial began). A minimum baseline accuracy of 80% on the 5-item, and 70% on the more difficult 7-item task was required to continue. If necessary, participants were allowed to repeat each of the training paradigms once and were disqualified if they failed to make the minimum accuracy score. Training and testing phase item strings had equal numbers of randomized and counterbalanced true positive and true negative probes.

4.2.2 Repetitive Transcranial Magnetic Stimulation

Increasingly, transcranial magnetic stimulation (TMS) is used as a research tool for mapping brain function and modification of neural processes during working memory tasks that engage PFC. TMS has been shown to be a safe, noninvasive means of stimulating the awake and alert human cerebral cortex brain in carefully screened subjects [159]. For instance, Koch et al. [81] used a repetitive TMS (rTMS) approach to disambiguate the spatial distribution and reciprocal interactions of different regions of the parietofrontal network in healthy human participants performing a spatial working memory task. These researchers compared the effect of 25 Hz rTMS on neural activity within PFC and posterior parietal cortex (PPC) during the delay and decision phases of the task. Trains of rTMS at 25 Hz were applied to PPC, premotor cortex, and right DLPFC alternatively during the two phases. They found that TMS during the delay phase in both parietal cortex and DLPFC interfered with performance. When 25Hz rTMS trains were applied during the decision

phase, they observed interference in DLPFC, but not in PPC or premotor cortex. The researchers interpreted that their results suggest two distinct neural circuits may exist in DLPFC: A local neural network subserving decisional processes and a second neural population functionally interconnected with PPC that was activated when spatial information was maintained temporarily in working memory [81].

Brandt et al. investigated the role of DLPFC and PPC in a visuospatial delayed response task [24]. Using high frequency rTMS, the researchers interfered temporarily with cortical activity in DLPFC and PPC during the maintenance period in order to determine whether rTMS to DLPFC or PPC during maintenance affects accuracy of memory guided saccades. Stimulation over DLPFC significantly impaired accuracy of amplitude and direction of saccades. They concluded that, within this network, DLPFC mediates the mnemonic representation and PPC with the sensory representation of spatially defined perceptual information [24].

Studies of the chronometrics of working memory phase evolution have used TMS in conjunction with fMRI to induce a temporary lesion that is intended to interfere with task performance, hence, identifying a causal role for the targeted cortical region in mediating working memory task performance. Such studies have shown that the effects of rTMS on cognitive performance are dependent on domain, experimental design, cortical target, stimulation frequency, timing, and duration of TMS application relative to the stage of task processing. Consequently, most studies of working memory and the role of PFC have been aimed to identify whether and precisely when functional disruption of PFC would degrade working memory performance. For example, Mull and Seyal aimed to determine whether transient functional disruption of DLPFC would impair performance in a working memory task in which participants were shown sequences of letters and asked to decide if the letter just displayed was the same as the letter presented three trials back [106]. Single pulse TMS was applied over DLPFC between letter presentations. TMS to left DLPFC caused increased errors relative to controls. The authors concluded that their results supported the idea that DLPFC played a crucial role in this paradigm. Deleterious effects of TMS to LPFC on working memory were also observed by Osaka et al. in a study using paired pulse TMS [111]. Albeit fewer in number, studies showing beneficial effects of TMS on working memory have also observed a critical role for stimulation frequency and timing relative to the phase of task evolution. For example, individual power and peak frequency attributes of alpha band oscillations have been shown to predict performance on a working memory task [79, 94]. In a previous study, the present authors designed a combined EEG–TMS study of the interaction between working memory Ne, alpha band oscillations, and 10 Hz rTMS applied to DLPFC prior to the onset of a Sternberg task [7]. We predicted first that compared with sham 10 Hz rTMS would significantly reduce psychomotor

processing speed indexed as RT. Second, we predicted that improvement in RT would be associated with a 10Hz rTMS induced increase in pretask alpha power and pretask alpha band phase synchrony. We identified a robust association between 10 Hz rTMS perturbed individual mean peak alpha frequency and alpha spectral phase synchrony on task performance.

4.3 Experimental Methods

4.3.1 Experimental Design

This was a single blind, sham controlled study (Figure 4.3). In 32 healthy participants, active or sham 10 Hz rTMS was directed to either right or left DLPFC 10 sec prior to the onset of a Sternberg short term verbal recognition memory paradigm. Five 5-sec trains of 10 Hz rTMS (for a total of 250 pulses) were applied to DLPFC ending 10 sec prior to task onset. Stimulation was followed by 48 single trials of 6- and 8-item strings of capital letters that were randomized and counterbalanced. Next, 250 additional pulses of 10 Hz rTMS were delivered to DLPFC of the same side, followed by a second set of 48 single trials of 6- and 8-item strings of capital letters. After a 15- to 20-min rest, the same sequence of events was repeated on side 2. Outcome measures were pre- and poststimulation accuracy, reaction times, and neural efficiency by stimulation group.

4.3.2 Participants

The Investigational Review Boards for the University of Utah and Veterans Health Affairs of the Salt Lake City Health Care System approved this study. Thirty-four medically healthy, right handed participants, ages 18–55 years, from the University of Utah and the Salt Lake City community were admitted to the study free of exclusionary conditions for TMS [159] and without a past or present psychiatric disorder. All participants had a negative screen for drugs of abuse on the morning of the study. Participants were instructed not to drink caffeinated beverages for 12 hours prior to testing. Participants were compensated for their time. Group assignment to stimulation (active or sham) and side first stimulated (S-1) were pseudorandomized and counterbalanced.

4.3.3 Paradigm Training

Task stimuli and computation of reaction times and accuracy were derived using Presentation 9.90 software (Neurobehavioral Systems, San Francisco, CA). A single trial of this task was composed of four phases: stimulus (a 5- or 7-item string of uppercase consonants and vowels was presented at the center of a 19-inch video monitor for 1.5 sec), delay (participants focused on a fixation marker appearing at the center of the computer monitor while maintaining the previous string in memory for 1.5 sec), probe (a single uppercase letter appeared at center screen for 3 sec), and response (participants were instructed to decide as quickly and accurately as possible
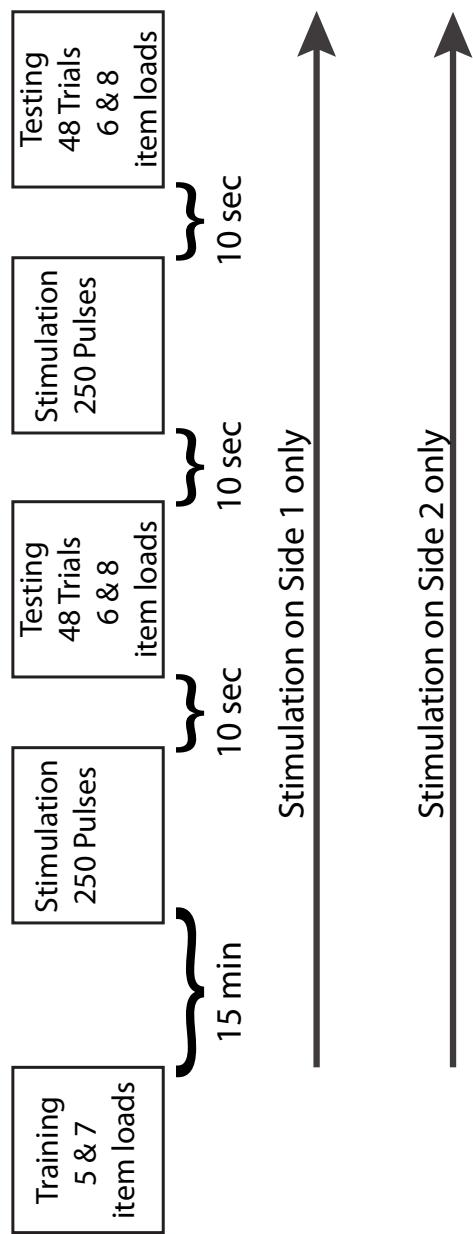
Figure 4.3: Illustration of the experiment performed using rTMS to manipulate working memory. Participants were trained using lower load examples. After training, each participant is stimulated on side 1, and then tested with 48 higher load instances of the Sternberg Paradigm. They were then stimulated on side 2 and presented an additional 48 trials.

whether the probe was present in the previous string, and to respond by pressing one of two keypad buttons; if there was a response within 3 sec, it was recorded, and the next trial began. If there was no response, the trial was marked as an error, and the next trial began). A minimum baseline accuracy score of 80% on the 5-item, and 70% on the more difficult 7-item task was required to continue. If necessary, participants were allowed to repeat each of the training paradigms once and were disqualified if they failed to make the minimum accuracy score. Training and testing phase item strings had equal numbers of randomized and counterbalanced true positive and true negative probes.

Participants were trained and tested between 11:00 a.m. and 1:00 p.m. They were familiarized with the task and response procedures by practicing to use their right first finger to press one of two buttons (YES or NO) on a handheld keypad as comfortably quickly as possible while attending to a fixation cue at the center of a 19-inch computer monitor. Baseline values of accuracy and reaction time were derived from 20 randomized and counterbalanced 5- and 7-item single trials of the Sternberg task.

4.3.4   Subject Stimulation and Testing

During the experiment, participants were comfortably seated in an individually adjustable chair. Participants single pulse TMS resting motor threshold was determined by identification of the lowest level of TMS machine output that would induce a visible contraction of the first dorsal interosseous muscle of the dominant hand on 5 of 10 pulses delivered with a 10-sec interpulse interval. A MagStimTM Super Rapid stimulator with physically identical sham and active air cooled figure-of-eight coils (Magstim, Whitland, Dyfed UK) was used to deliver TMS pulses directed to either left or right DLPFC. The sound generated by capacitor discharge was identical in the two coils; the sham coil did not generate a sensorimotor stimulus. A custom designed mechanical coil holder was used to maintain contact between the midpoint of the 70 mm figure-of-eight coil and the location of the frontal scalp overlying right or left DLPFC. To determine the scalp correlates for targeting of DLPFC, we used the measures 1020 EEG system [68]. Positioning of the coil on the scalp was determined by the measured position of EEG sensors F3 (left) and F4 (right). Participants used disposable in ear sound protection during the experiment. Side stimulated first (S-1) was randomly assigned and counterbalanced.

A brief period of acclimation to the laboratory environment was followed by the onset of five 5-sec trains of 10 Hz rTMS with an intertrain interval of 10 sec at 100% of resting motor threshold for a total of 250 pulses of 10 Hz rTMS. Stimulation ended 10 sec prior to the onset of the first 48 single trials of the task. In order to minimize the effects of practice on performance, single trials in the testing phase were composed of more difficult 6- or 8-item uppercase letters pseudorandomly

presented and counterbalanced for true positive and true negative probes. This sequence was repeated once more on S-1. Following a 15- to 20-min rest period, the entire sequence was repeated on S-2; hence, the study totaled 192 single trials and 1000 pulses of 10 Hz rTMS.

4.4   Experimental Results

Thirty-four right handed men and 12 women recruited from the Salt Lake City, Utah community participated in this study. Two subjects in the actively stimulated group were unable to complete the study due to the development of moderate muscle contraction headaches. In both cases, the headaches spontaneously resolved. There were no adverse events in the sham group. All participants achieved the minimum 5- and 7-item training accuracy on their first attempt.

4.4.1   Analyses

Skewness and kurtosis in the distribution of dependent variables were identified and normalized with square root transformation. Repeated measures ANOVA with Bonferroni correction for multiple comparisons was used to identify the main effects of stimulation, S-1, and interactions between stimulation and S-1. Dependent variables were covaried with resting motor threshold and age. Pairwise comparisons were used to determine the within subject effects of active and sham 10 Hz rTMS on accuracy and reaction time by task load.

4.4.2   Demographics

There was no difference in mean age between stimulation groups or by sex. Mean age by Stimulation group: active = 28.6 (σ = 5.68); sham = 30.88 (σ = 11.2); independent samples two tailed t test for equality of means; equal variance not assumed: t(31) = 0.707, p = .478. Mean age by Sex: male = 28.75 years; female = 29.50 years; two tailed t test: t(30) = 0.25, p = .80. There was equal distribution of the sexes between groups [Pearson uncorrected chi-square test: χ^2(1, n=32)=0.139, p=.710].

4.4.3   Resting Motor Threshold

The groups were evenly matched for the intensity of stimulation indexed as resting motor threshold and expressed as a percent of maximum machine output [active resting motor threshold = 67%, sham = 66%; two tailed t(34) = 0.252, p = .4484].

4.4.4   Statistical Analysis

There was no main effect of S-1, and no interaction between S-1 and stimulation. Measures of dependent variables did not covary with resting motor threshold, sex, or age. In the absence

of a main effect of S-1, prestimulation and poststimulation values of dependent variables were collapsed across side for subsequent analysis. There were no prestimulation differences between groups regarding accuracy, reaction time, or neural efficiency. There was, however, a robust main effect of stimulation on poststimulation reaction time and neural efficiency, but not accuracy, as presented in Tables 4.2 and 4.3.

In this study, 10 Hz rTMS was directed to right or left DLPFC 10 sec prior to task onset, predicting there would be a main effect of stimulation on accuracy, reaction time, and neural efficiency. We observed a robust main effect of stimulation in the comparison of poststimulation reaction time and neural efficiency, but not accuracy. The mean poststimulation reaction time in the active group was 219 msec (± 0.16) faster than the prestimulation baseline, whereas the mean poststimulation reaction time in the sham group was 30 msec (± 0.16) slower than the mean baseline reaction time. Pairwise comparisons of pre- and poststimulation reaction time, accuracy, and neural efficiency were computed by comparing the relatively easier 5- and 7-item (prestimulation) measures with the 6- and 8-item (poststimulation) measure. Pairwise comparisons of the 5- and 7-item training reaction time data (Table 4.4) show a small but nonsignificant within subject difference in both groups for this comparison. Specifically, pairwise comparisons in the actively treated group showed that compared with the larger 6-item and 8-item poststimulation loads, the 5-item prestimulation reaction time was slower when the converse would be expected. Similar results were seen when the comparison of the prestimulation 7-item reaction time was compared with the poststimulation 8-item reaction time. In the sham group, the same comparisons failed to show a difference for pre- or postmeasures of reaction time, accuracy, or neural efficiency (Figures 4.4 and 4.5).

4.5 Visualizing Working Memory

Signals and time series represent one of the most ubiquitous forms of data in science and engineering. With the advent of electronic equipment, time series acquisition and analysis became commonplace in many fields of research.

The analysis of time series data relies on the detection of both the obvious and the subtle spectral properties of signals. More apparent aspects of signals, such as periodicity and general patterns and trends, are often determined by direct inspection of the time series itself. Weber et al. revived a method of time series visualization first proposed by Antonio Gabglio in 1888 whereby the input data were visualized on an expanding spiral [56, 161]. This method allows periodicity of a signal to be established visually. Exploring recurrences, patterns, and periodicity while preserving the temporal locality of each occurence was discussed in arc diagrams by Wattenberg [160].

Examining time series data to extract patterns, regardless of periodicity, is another important aspect of signal processing. Of particular interest is the occurrence of a given pattern or anomaly in

Table 4.2: Effects of 10 Hz rTMS on the accuracy, reaction time, and neural efficiency of working memory.

Descriptive Statistics				
	Active		*Sham*	
	Mean	SD	Mean	SD
Pre rTMS Accuracy	0.920	0.07	0.970	0.03
Post rTMS Accuracy	0.870	0.06	0.860	0.05
Pre rTMS Reaction Time	1.117	0.08	1.010	0.10
Post rTMS Reaction Time	0.874	0.05	1.014	0.11
Pre rTMS Efficiency	0.832	0.17	0.969	0.21
Post rTMS Efficiency	0.995	0.13	0.866	0.20
Mean Efficiency Change	1.200	0.30	0.90	0.10

Table 4.3: Significance analysis of the effects of 10 Hz rTMS on the accuracy, reaction time, and neural efficiency of working memory.

Repeated Measures: ANOVA		
	$F(df = 1, 30)$	*Signficance*
Pre rTMS Accuracy	0.019	0.892
Post rTMS Accuracy	0.012	0.915
Pre rTMS Reaction Time	3.170	0.086
Post rTMS Reaction Time	7.380	0.011
Pre rTMS Efficiency	1.628	0.212
Post rTMS Efficiency	5.344	0.028
Mean Efficiency Change	14.253	0.001

Table 4.4: Pairwise comparisons of the effects of 10 Hz rTMS on accuracy, reaction time, and neural efficiency.

Two-tailed Pairwise Comparisons		
	t	p
Active		
Pre-vs-Post Accuracy	2.362	0.0340
Pre-vs-Post Reaction Time	4.417	0.0010
Pre-vs-Post Neural Efficiency	-4.993	0.0002
Sham		
Pre-vs-Post Accuracy	3.113	0.008
Pre-vs-Post Reaction Time	-1.264	0.228
Pre-vs-Post Neural Efficiency	0.986	0.342
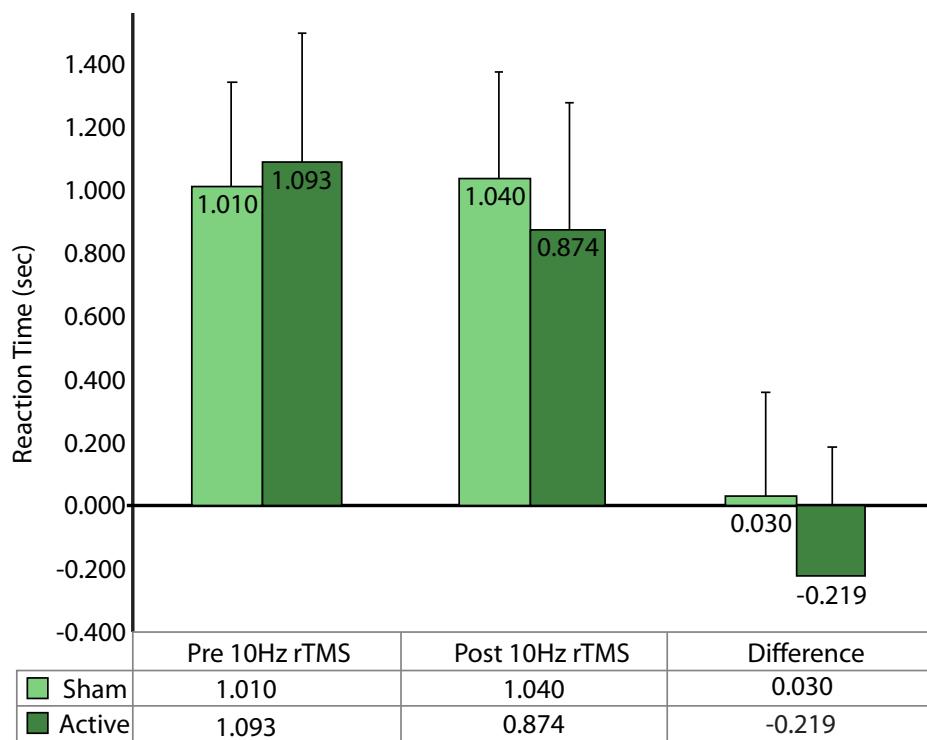
Figure 4.4: Comparative effects of active and sham 10 Hz rTMS on working memory reaction time. Compared with sham, the group treated with active 10 Hz rTMS applied to the dorsolateral prefrontal cortex 10 sec. before onset experienced a 219ms enhancement of reaction time.
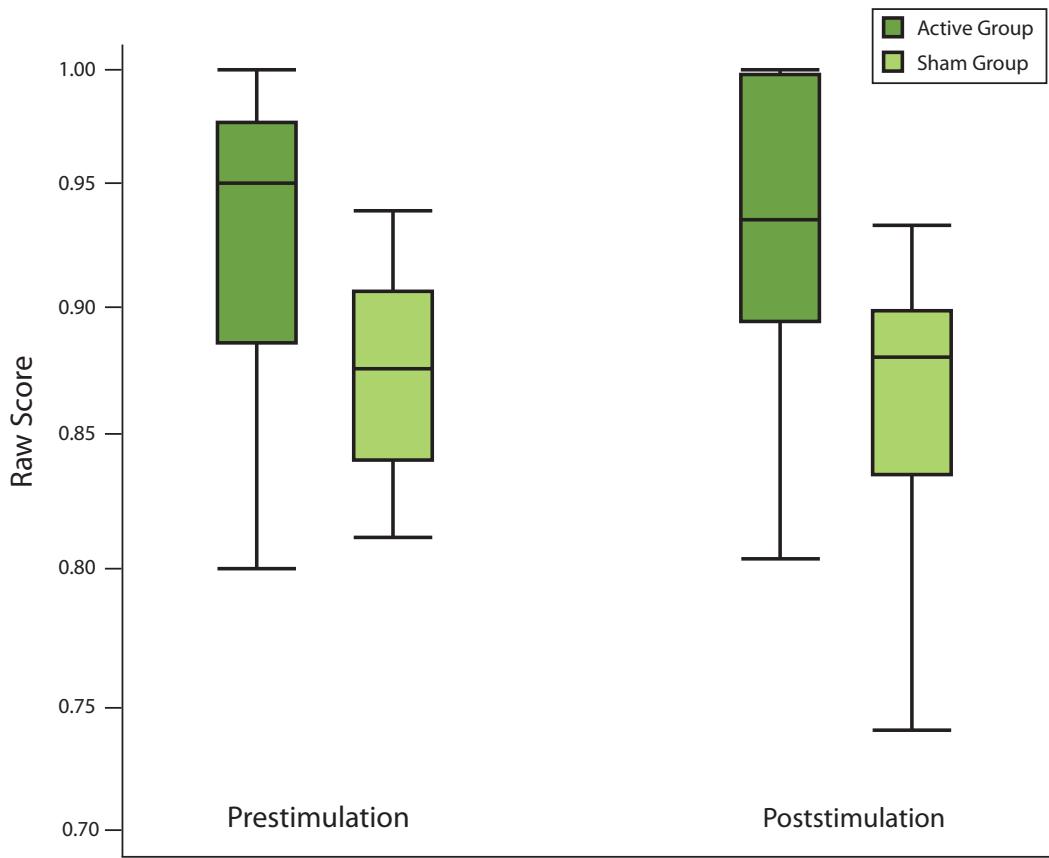
Figure 4.5: Comparative effects of active and sham 10 Hz rTMS on Sternberg task accuracy. There was no signficant change in accuracy after application of 10 Hz rTMS in either stimulation group.

more than one time series. Yankov et al. describe a system in which dot plots are used to assess and visualize the similarities and differences of two signals [164]. Likewise, Lin et al. demonstrated that symbolic representations of time series can be used to detect, extract, and visualize nontrivial patterns found across many distinct signals [91, 92].

When processing collections of signals, it is important to align the various time series to ensure that the acquisition times correlate correctly. This alignment is essential for the proper extraction of features in the signal. To alleviate the problems associated with acquisition timing errors, Chu et al. developed time warping schemes to properly align these data [34]. Additionally, Aach et al. present a detailed discussion of various time warping algorithms as they are applied to the time series based evaluation of RNA expression collected for large numbers of genes [2].

Visualization of raw time series data is effective in extracting overt qualities of the input data. However, subtle features, such as estimated power spectral density, require a different class of processing techniques. Although the theory of frequency analysis for time series had been formalized in the 19th century, its broad application was not realized until relatively recently [23].

The standard frequency analysis techniques, Fourier and Fast Fourier Transforms, treat entire signals as atomic elements. This type of analysis loses the time dependent aspect of evolving time series. To address this, the Short Time Fourier Transform (STFT) is often used to produce a time frequency plane representing the frequency content of a signal as it evolves in time [5]. Another decomposition that highlights time frequency evolution is the Wavelet Transform. Stéphane Mallat furthered hierarchical wavelet based techniques in engineering applications [96] enabling wavelet transforms to be applied to a wide range of disciplines and signal types [12, 101].

Aggregating related images into a single visualization is often an effective means of determining relationships between the various members of the input set. Design Galleries allow an intuitive means of exploring large parameter spaces by examining images organized by the parametrization used to generate them [97]. Dimensionality reduction techniques such as Locally Linear Embedding [126] and IsoMap [152] are solutions that preserve local neighborhood relationships. These methods often present data in an exploration interface similar to design galleries. Although these methods provide solutions to many of the problems encountered during the visualization of EEG data, no method exists to adequately visualize these data in their entirety; instead, each method focuses on a small subset of the dimensions inherent in the data.

4.5.1 Data Acquisition and Processing

Many techniques developed to inspect biological time series data have proven to be useful in various analysis tasks, but they do not focus on the spectral properties fundamental to the study of working memory. The applications of time frequency analysis and visualizations to biological

signals are particularly interesting. McNames et al. explored various visualization techniques applied to different classes of biologically generated time series [100]. With the introduction of highly accurate, high density electro- and magnetoencephalography arrays, a flood of new data regarding signals generated in the brain has made the simultaneous visualization of many time series a necessity. A large body of work has focused on the processing and visualization of multiple signals acquired through EEG [150, 151].

In order to properly study the function of working memory in the human brain, data in the form of EEG time series are acquired, processed, analyzed, and interpreted. Unfortunately, while several visualization techniques are applicable to EEG data, a more specific set of requirements is imposed by the study of spectral dynamics in working memory. One requirement, the ability to visualize many time frequency planes simultaneously, precludes the use of many common EEG visualization methods from being employed. While these methods are both practical and useful in the study of EEG data as a whole, they do not provide an adequate solution to the specific needs of this study.

Our study of working memory began with the collection of 64 channel EEG data sampled at 1 kHz during a working memory task. Each participant in the study was assigned to either the control or experimental group to determine the rTMS stimulation parameters. The working memory task was administered using a 142 single trial Sternberg paradigm [145] after the application of rTMS directed at the DLPFC, the brain area thought to be primarily responsible for working memory function. The resulting dataset is a collection of 64 individual signals (one signal for each EEG electrode) with labels segmenting the signals into each of the 142 single trials. Combined with sensor positions, each EEG signal can be reliably mapped to the spatial location from which it was collected.

4.5.2   Data Analysis

After acquisition, each segmented signal must be processed and analyzed. We examine the spectral evolution of the alpha frequencies. These analyses require the use of time frequency decompositions to extract the spectral content from the EEG data. Each decomposition maps a one-dimensional, temporal input signal to a two-dimensional domain of time and frequency. In this work, we take advantage of the various properties of the S-Transform [147]:

$$S(\tau, f) = \frac{|f|}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-\frac{(t-\tau)^2 f^2}{2}} e^{-i2\pi f t} dt,$$

where $h(t)$ is the input signal, and f and τ are the frequency and time point being calculated, respectively.

While there exist several transforms capable of representing a signal in the time frequency domain, each transform brings with it a unique set of advantages and restrictions. The S-Transform is similar to the generalized Wavelet Transform [96], as it represents a similar packing of the time frequency plane with basis atoms. An important consequence of the S-Transform's choice of basis is a direct mapping to the Fourier Spectrum. In contrast, both the Continuous and Discrete Wavelet Transforms represent frequency in terms of the dyadic dilation of the basis atom — the *pseudofrequency*. Figure 4.6 illustrates the decomposition of a signal in time to a time frequency representation of energy density. This time frequency plane shows alpha activity in the parietal cortex evolving over time.

The notion of frequency in terms of the Fourier Spectrum is important for the study of working memory. The Fourier Domain forms a well understood base for analysis, whereas wavelet based methods rely on more abstract mappings to frequency. Additionally, the evolution of the frequency spectrum over time is a key component in the study of brain mechanics. The inclusion of the S-Transform into our processing pipeline was motivated by its ability to provide a direct mapping to the Fourier Domain while maintaining an adaptive resolution scheme like the Wavelet Transform. As with any time frequency representation, the image like nature of the two-dimensional domain complicates the simultaneous visualization of a large collection of time frequency planes. The expression of a dataset of this type as a whole is important when exploring the correlations and interdependencies between the various signal sources comprising the data.

4.5.3   Data Visualization

Because the time frequency data collected and derived from EEG signals are multidimensional, no single visualization technique can be employed to examine every aspect of them. Topographic maps show EEG data collected on the scalp by coloring a surrogate head according to the voltage identified at that time and position [82]. This organization allows topographic maps to present a view of EEG data that reinforces the spatial organization given by the physical sensor locations. Instead of overlaying multiple EEG traces to accommodate many sensors to show general trends in collected signals, topographic maps like those shown in Figure 4.7 present two-dimensional views of the data. Viewing the data in this manner enhances the spatial information associated with alpha band activation across the entire cortical area covered by the EEG sensors.

Typically, topographic maps display information associated with signal voltages acquired from EEG sensors. However, Figure 4.7 shows a generalization of this technique using scalars associated with spectral power at the alpha frequencies. This is done to present spatial information associated with the given frequency band. Although topographic maps enable scientists to visualize the spatial relationships of EEG signals acquired on the scalp, they can only be formed with respect to a single
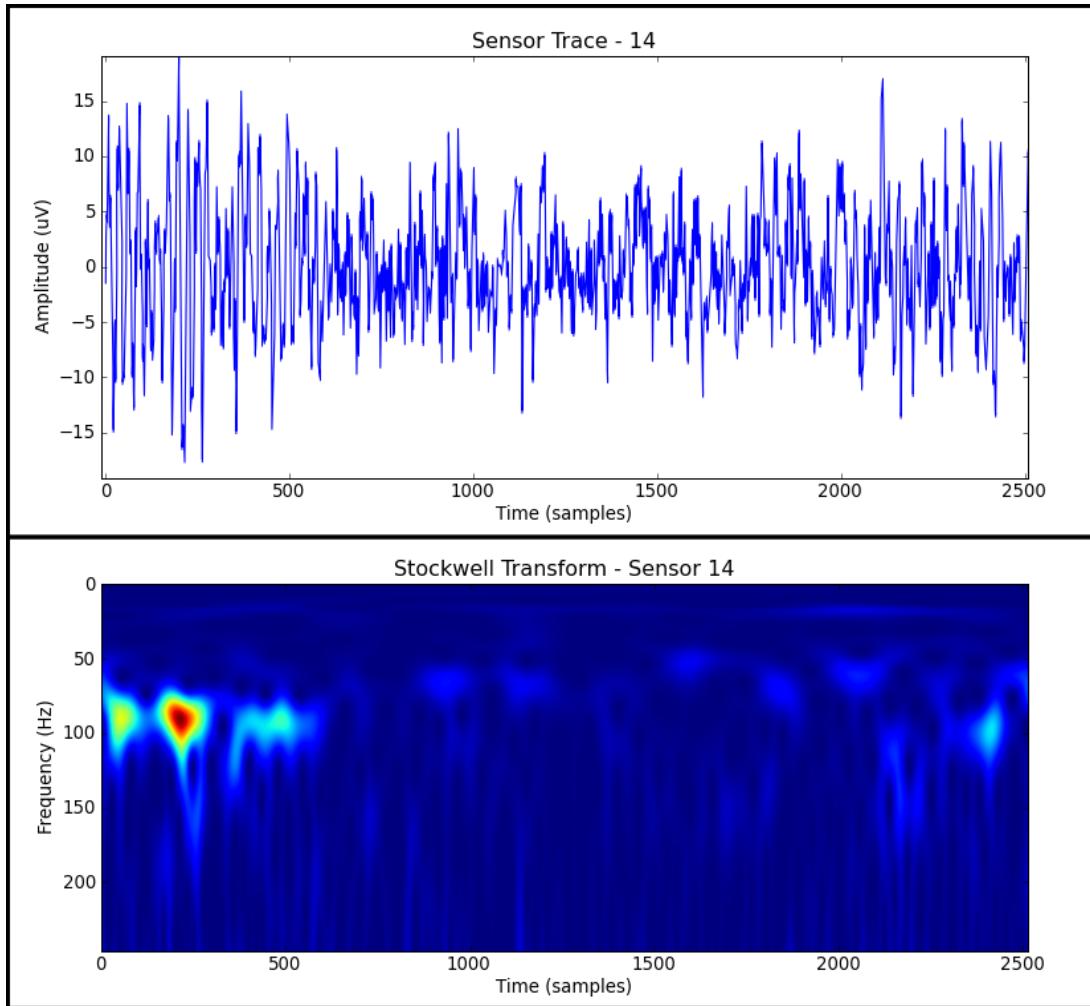
Figure 4.6: The Stockwell Transform of a single sensor's data. (Top) A plot of the raw data trace. (Bottom) The S-Transformed representation of the sensor's raw data. Since this visualization shows all representable frequencies at each timestep, this view allows the time series' frequency evolution to be thoroughly analyzed.
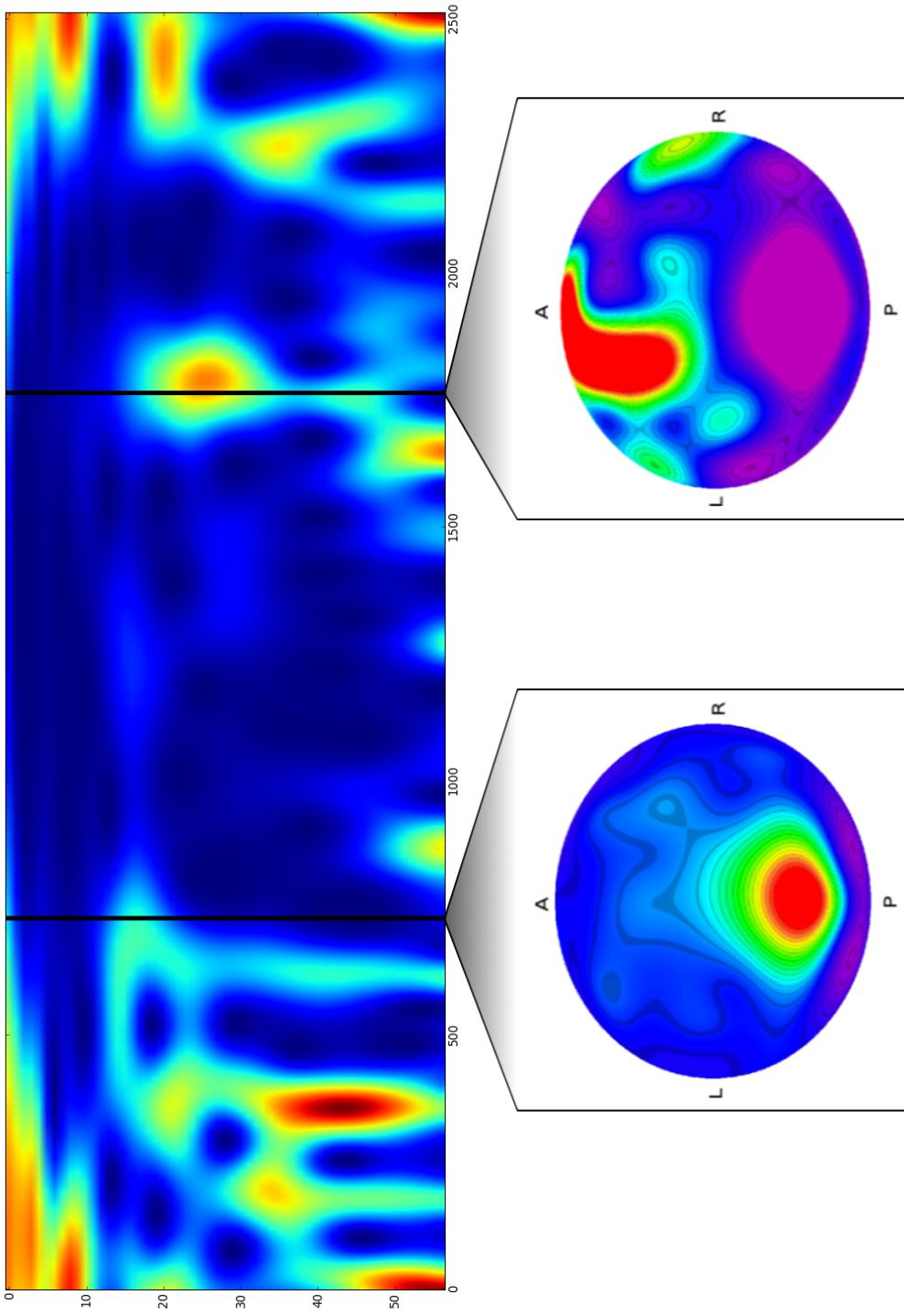
Figure 4.7: Topomaps as seen in the lower half of this figure depict the activity of all sensors in an EEG network at a single time and frequency. In this case, the topomaps represent the activity of alpha spectral power throughout the brain. This view of the data excels at representing spatial relationships of a frequency at a single point in time, but can visualize neither mutiple frequencies simultaneously, nor their evolution over time.

time and scalar value. Animation of topographic maps helps visualize time dependent data, but an overview of the spectral dynamics is lost. This limitation inhibits the visualization of spectral properties as they evolve in time and frequency.

Unline topographic maps that use individual samples from each time frequency plane, multidimensional scaling techniques, such as IsoMap, [152] produce groupings of time frequency planes using each plane's entire content. Figure 4.8 displays the result IsoMap's grouping using our distance measure described in Section 4.5.4. The figure demonstrates that IsoMap's pair wise distance comparison provides a mechanism that appropriately clusters time frequency planes with similar features. While IsoMap provides a representation of all planes in relation to one another it is difficult to visualize all time frequency planes simultaneously using this view of the dataset. We adopt the notion of distance measures for grouping spectral planes to construct time frequency volumes.

4.5.4   Constructing Time Frequency Volumes

Because EEG collected at the scalp measures voltage generated by a relatively large portion of the cortex, neighboring sensors often exhibit strong correlations in the acquired time series. However, as processing tasks migrate to different parts of the brain, these correlations strengthen and weaken accordingly. Since working memory is identified by recognized spectral patterns, scientists can measure the overall performance of a person's working memory [7].

Exploring a dataset in terms of the correlations exhibited between its individual components remains an effective way to explore and analyze data [155]. Gonzalez and Woods show that correlation provides an accepted method for measuring the similarity between two different images [63]. In the case of this working memory study, correlations are found between time frequency images collected by the various sensors and the resulting decomposition of the given signal. We use the correlation between these image planes to measure similarity between the time frequency decompositions we calculate.

We analyze working memory through correlations between the various sensors in the EEG network and the spectral patterns associated with its activation. Since the degree of correlation and its time lag are both important in determining links between areas of the brain, we must respect these values when determining similarities between time frequency planes.

The degree of correlation between images f and g is given by the global maximum of the convolution image:

$$c = max(f \star g).$$
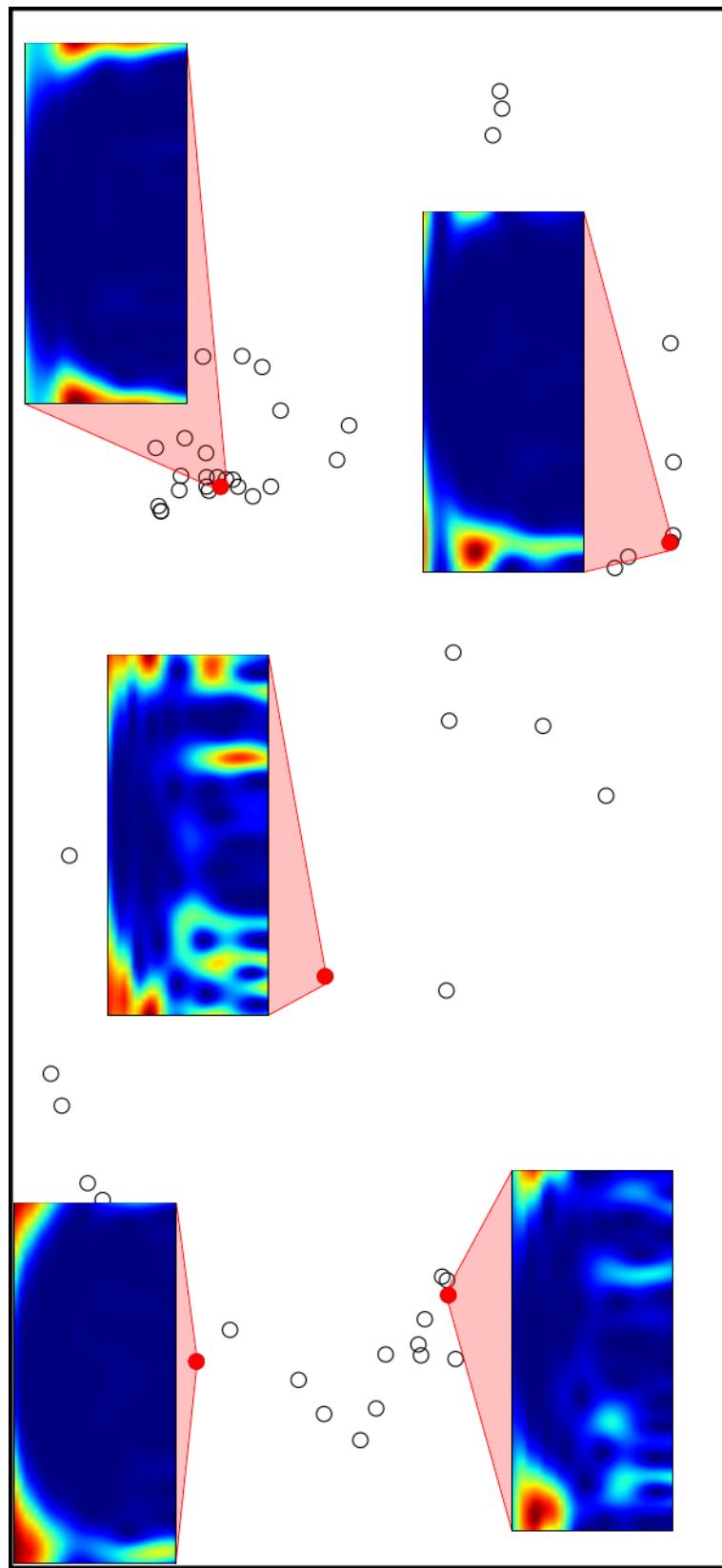
Figure 4.8: Processing all time frequency planes using the **IsoMap** algorithm shows possible neighborhoods induced by our distance measure. While brushing eases the task of exploring each plane's neighborhood, examining closely related time frequency representations is difficult without displaying a prohibitive number of images simultaneously.

We combine the degree of correlation with the euclidean distance between the center of the time frequency plane, $o(f)$, and the location of the maximum correlation to form a weighting coefficient $c'_{x,y}$:

$$c'_{x,y} = \|o(f) - argmax(f \star g)\|_2,$$

where x and y represent the coordinates of the global maximum of $f \star g$. Since the analysis of time frequency evolution of working memory places more importance on frequency localization than time localization, a two-dimensional gaussian weighting function is used to independently control any shifts in the correlation separately in the time and frequency domains:

$$w(x,y) = e^{-(\frac{(x-o_x(f))^2}{2\sigma_t^2} + \frac{(y-o_y(f))^2}{2\sigma_f^2})},$$

where σ_x and σ_y are the gaussian parameters in each dimension and $o_x(f)$ and $o_y(f)$ are the x and y components of the offset of the global maximum and the center of the time frequency plane. This Gaussian is then combined with the previous weighting coefficients to yield the distance measure:

$$\|f,g\|_2 = 1 - c'_{x,y} \cdot c \cdot w(x,y).$$

This formulation incorporates the time lag of correlations, enhancing the analysis of spatial oragnization of neural circuitry.

This distance measure describes the normalized similarity between two time frequency representations, f and g. By incorporating the maximum correlation's distance from the center of the image, changes corresponding to shifts in both the time and frequency domains are independently weighted. This independence is important to the study of brain mechanics as spectral properties appearing at different times indicate the order in which various areas of the brain are activated by a single stimulus. Shifts in the frequency domain are suppressed by narrowing the gaussian weighting function in the frequency domain. This weighting prevents frequencies generally indicative of processing outside the context of working memory from influencing the ordering.

4.5.4.1   Ordering the Time Frequency Volume

The function ORDER-VOLUME describes the greedy process governing the pair wise comparisons used to form the correlated signal volume. The ORDER-VOLUME method takes a collection of time frequency planes, C, and a user defined key plane, k as inputs to form the final ordering of the correlated volume, C_v. As the ordering is generated by comparing pair wise distances across the

time frequency collection similarly to the comparisons performed in IsoMap, the resultant volume retains neighborhoods of similar planes.

ORDER-VOLUME(C, k)

```
1   output = List()
2   while size(output) < size(C)
3       do
4           V_c = List()
5           for each C_i in C
6               do
7                   cor = correlate(C_i, k)
8                   V_c.append(cor)
9
10          output.push(V_c.max())
11          V_c.remove(V_c.max())
12          k = V_c.max()
13
14  return output
```

The ORDER-VOLUME method is responsible for computing all pair wise correlations with the user specified key plane. This operation is equivalent to traversing a fully connected graph by locally maximizing path lengths defined by our distance measure. By placing the highest correlated plane in the output volume and using it as a new key plane to measure similarities with the remaining planes, a greedy ordering is induced that highlights common features in a local setting while also forming a global ordering separating time frequency planes that are substantially different from the key plane. In this way, the ORDER-VOLUME method forms an approximation to a multidimensional scaling solution.

Figure 4.9 demonstrates how the appropriate placement of individual slices within a volume affects the correlation between neighboring planes. In this figure, a synthetic dataset was reordered based on the global correlation with a key plane as well as our ORDER-VOLUME method. Since the ORDER-VOLUME method of progressive correlation greedily places highly correlated planes into the volume, it has a greater degree of interslice correlation than does this global method.

4.5.4.2   Correlation and Coherence

Because the greedy approach to ordering by correlation described above is sensitive to user specified input, different orderings of the same time frequency collection are generated for different user defined key planes. This flexible ordering of the volume is useful in exploring the spectral dynamics both across times and frequencies of interest as they change due rTMS application.
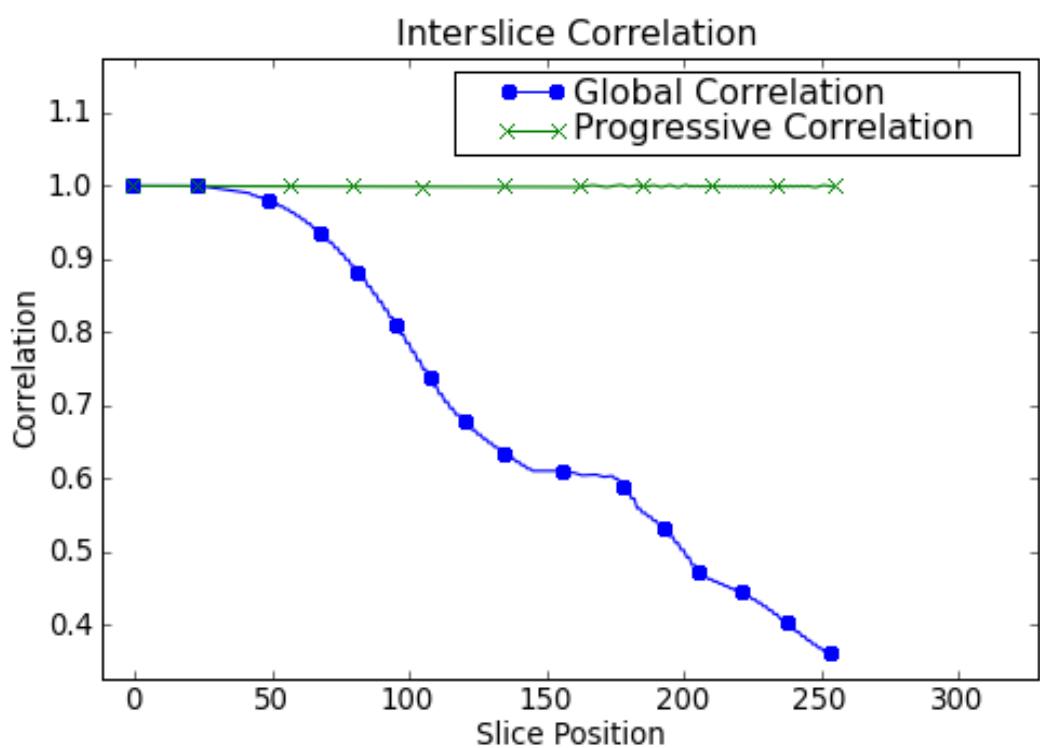
Figure 4.9: Interslice correlation depends greatly on the order the slices appear in the new volume. Ordering the slices of a synthetic dataset using a global correlation method does not maintain high correlation levels throughout the entire volume as our greedy progressive correlation.

Coherence between slices is important to maintain because as correlation measures the global similarity between two inputs, coherence measures the similarities between them on a feature-by-feature level in the Fourier Domain. The coherence, $r(x, y)$, of two signals is given by the normalized cross spectral density of the two inputs, where the cross spectral density is R_{xy} and the autospectral density is R_{xx} [32]:

$$r(x, y) = \frac{R_{xy}^2}{R_{xx}R_{yy}}$$

To measure the coherence of an ordered volume of time frequency planes, a pair wise measure of coherence is made. Coherence and correlation are similar metrics. The greedy approach presented here maintains high levels of coherence between slices of the volume as compared to other methods. This coherence allows scientists interpreting the volume to track specific activation patterns as the brain activity moves throughout the cortex.

While this technique is well suited for visualizing features present in a dataset composed of related, yet independent members, it can also be used to enhance the coherence between planes in other volumetric data. Figure 4.10 depicts a synthetic dataset consisting of multiple, concentric spheres. Due to the spherical nature of the dataset, it is symmetric about any plane intersecting the center of the volume. Since symmetry implies that congruent slices of a volume are separated by a number of other slices, our greedy approach to coherence ordering will create volumes that are not symmetric with respect to the planes being analyzed. While symmetries in datasets pose interesting questions regarding the reconstruction of volumetric datasets such as in Figure 4.10, it is advantageous when aggregating related, yet independent, time frequency planes to know that symmetry in the resulting volume will be avoided. Avoiding symmetry in time frequency volumes allows features in time frequency planes to be ordered such that they maximize the coherence between them and their features.

4.6   Discussion

This study compared the relative effects of active and sham 10 Hz rTMS directed to the right and left DLPFC on psychomotor processing speed and accuracy, indices of neural efficiency, in healthy participants performing the Sternberg task, a short term verbal recognition memory paradigm that has been shown to model the association between working memory load and task behavior [146]. Previous studies of the role of DLPFC in working memory have been designed to identify the temporal and spatial dynamics of the computational role of DLPFC in a working memory network. For example, Osaka et al. investigated the role of executive function in human left DLPFC in eight normal participants using low frequency TMS after fMRI activation confirmed a role for

(a)

(b)
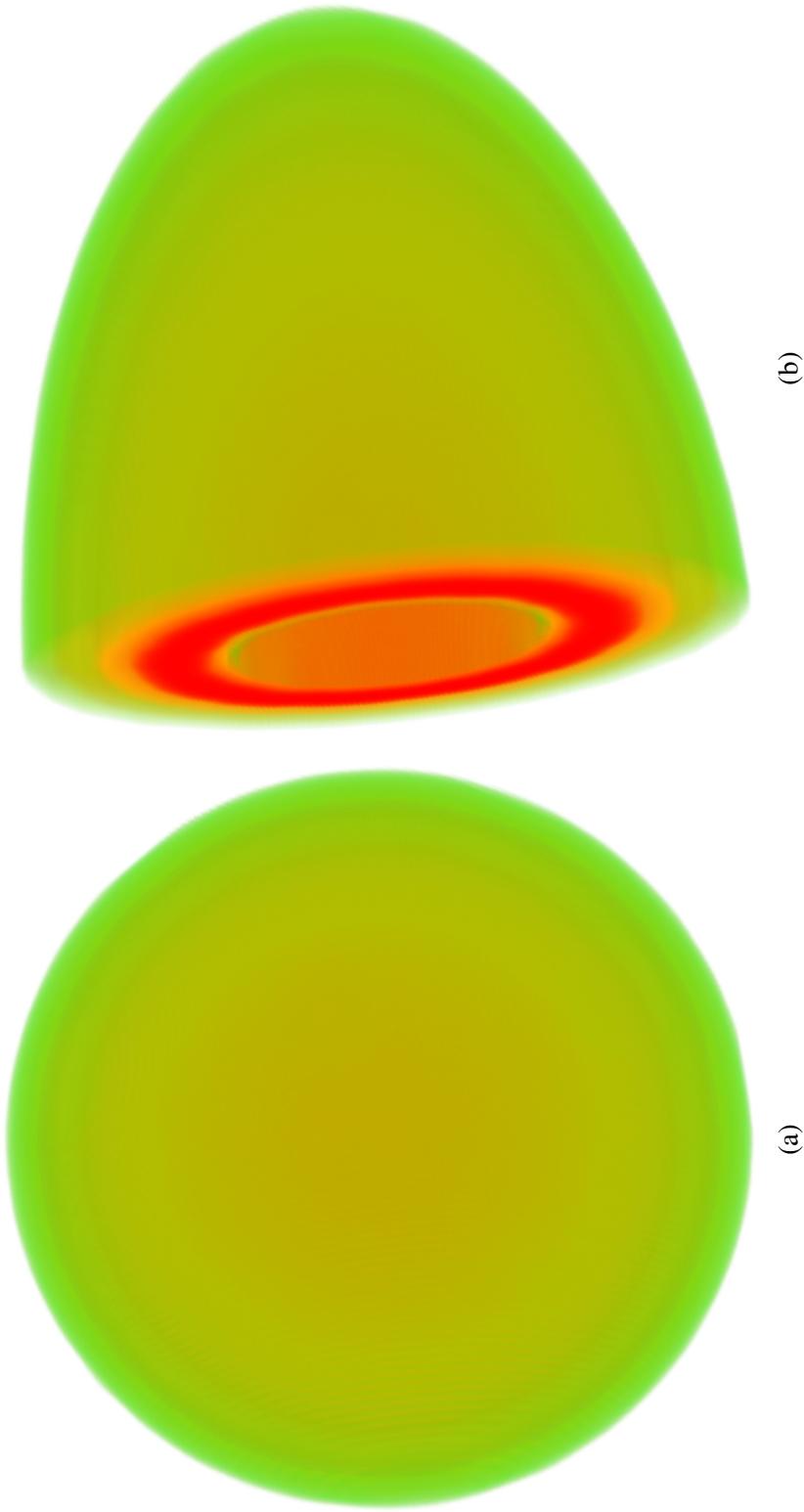
Figure 4.10: Volumetric datasets have high coherence between adjacent slices, giving rise to features that exist in all three dimensions of the volume. Here, a synthetic dataset composed of three concentric spheres (a) is shown. By reordering slices in this symmetric volume using our progressive correlation with the midvolume plane as the key slice (b), the symmetry is exploited to form a different volumetric dataset with high coherence between slices.

left DLPFC [111]. TMS was applied to left DLPFC immediately after subjects finished reading sentences from the reading span test, a measure of verbal working memory. They found a significant deterioration of performance in this working memory task, hence, supporting a crucial role for DLPFC in working memory. Koch et al. used an rTMS approach to investigate the temporal and spatial dynamics of the parietofrontal network in normal subjects performing a spatial working memory task [81]. They aimed to compare neural activity in the relevant areas during the delay and decision phases of the task. Trains of rTMS at 25 Hz were directed to PPC and right DLPFC during the two phases. Premotor cortex was used as a spatial control. They observed that TMS interfered with performance during the delay phase in the posterior parietal and DLPFC sites. There was no effect for the control site. When rTMS trains were applied during the retrieval phase, interference was observed only in DLPFC. The authors concluded that their study identified an example of parallel processing in the parietofrontal network of spatial working memory during the delay phase. The results of Koch et al. represent additional evidence of the critical role of DLPFC during both the delay and the retrieval phases of a working memory task. Mottaghy et al. observed significant interference with working memory reaction time after 180 msec of left PFC stimulation [103]. Results in a similar vein were observed by other researchers [104, 106, 117]. The present study was designed, in part, to identify a putative role for DLPFC in the neural efficiency of a working memory task. Accuracy declined to a similar extent in both groups, likely due to the effects of memory load, hence, supporting the conclusion of previous studies that also found no effect of TMS on working memory task accuracy [42, 94, 95]. In turn, such findings suggest that the effect of 10 Hz rTMS on reaction time when directed to DLPFC 10 sec prior to task onset mediates the enhancement of neural efficiency. Additionally, these data suggest that the retrieval phase of this working memory task may be the temporal boundary of the observed effects.

Results of the present study are in line with Rypma et al., who set out to identify the neural bases of interindividual differences in cognitive performance [128]. They had participants perform a simple speed of processing task during fMRI scanning. In certain prefrontal cortical regions (PFC), faster performers used fewer neural resources than slower performers, whereas in other PFC and parietal regions, they used more. These results were interpreted to suggest that a critical determinant of interindividual differences in cognitive performance was the efficiency of interactions between brain regions. In order to perform well, slower individuals may utilize more prefrontal resources than faster individuals [128].

Taking into account the lack of difference between groups in baseline reaction time (Table 4.5), together with a robust reaction time enhancement with active stimulation, and considering that participants in both groups had equal practice experience over the course of experiment, one might

Table 4.5: Pairwise comparisons of reaction time and memory load for each stimulation group. In the actively stimulated group, the 5-item prestimulation Reaction Time (RT) is slower than the 6- and 8-item variants; in the sham group, the 6- and 8-item reaction times are slower than the 5-item variants, as would be expected.

Group	*RT(I)*	*RT(J)*	*Mean Difference(I − J)*	*SEM*	*Significance*
Active	5-item	6-item	0.278	0.07	0.00610
		8-item	0.252	0.07	0.0170
	7-item	6-item	0.167	0.03	0.0001
		8-item	0.141	0.03	0.0005
Sham	5-item	6-item	-0.046	0.03	1.0000
		8-item	-0.101	0.04	0.1930
	7-item	6-item	0.098	0.04	0.1450
		8-item	0.052	0.05	1.0000

plausibly argue that our data support the conclusion that the poststimulation difference in task performance between groups is best explained as a difference between active and sham stimulation (Table 4.5). Even so, we recognize valid challenges to these conclusions. Our results were specific to accuracy, reaction time, and neural efficiency; even so it is possible that the effects were due to a priming effect of rTMS on psychomotor functions that are also mediated through prefrontal networks [124, 157]. Although we would suggest that it is unlikely the observed effects are strictly due to learning, inasmuch as both the sham and active rTMS groups had equal opportunity to practice, we recognize there may have been an interaction between 10 Hz rTMS and the learning of a response strategy that results in faster but not more accurate responses. An additional challenge to our conclusions is related to the capacity limits of short term memory [39, 98]. To overcome such limitations, study subjects may devise strategies, such as chunking to successfully perform at supracapacity limits, which can affect neural activity in a complex way, leading to potential complications in the interpretation of results. In view of the high capacity nature of the task we used in this study, it is possible that chunking or other mnemonic strategies may have been used to manage supracapacity aspects of task performance.

The underlying neurobiological mechanisms subserving the effects of TMS on working memory neural efficiency are unclear. In a previous exploratory study, the present authors found that 10 Hz rTMS applied to DLPFC in healthy humans resulted in a significantly enhanced measure of individual mean peak alpha frequency, and enhanced within band 1012 Hz phase synchrony [7]. These results are in line with Klimesch et al., who showed that mean peak alpha spectral frequency (individual alpha frequency) is related to memory performance [79]. Klimesch et al. showed that compared with attentional demands, memory performance exerted the strongest effect on individual

alpha frequency. The difference in individual alpha frequency between good and bad memory performers peaked when subjects retrieved information from memory. During retrieval, the individual alpha frequency of good performers was 1.25 Hz higher than that of bad performers.

Brignani et al. found that low frequency TMS over primary motor cortex induced a synchronization of the background oscillatory activity in the stimulated region [26], hence, suggesting that TMS effects may involve short term modification of the neural circuitry sustaining motor behavior [26, 77, 89]. Emergent evidence suggests that a diminished capacity to synchronize distributed neural assemblies mediating working memory may be a critical and enduring underlying mechanism of impaired working memory in a broad array of seemingly disparate neuropsychiatric disorders such as schizophrenia [10, 51, 55], depression [93], Alzheimers disease [112, 165], and autism [156].

Despite such recognition, there is no treatment for impaired memory in any such disorder. TMS has been shown to mediate events at the neural synapse [59]. In light of previous work, this suggests that such effects may be related to 10 Hz rTMS induced phase state perturbation [7]; if so, it may also suggest a role for TMS as a clinical tool for neural circuit repair or enhancement [116, 135]. In a study with 24 subjects, Hamidiet al. used high frequency rTMS to evaluate the role of DLPFC in memory guided response to two different types of spatial working memory tasks [66]: one requiring a recognition decision about a probe stimulus (operationalized with a yes/no button press), and another requiring direct recall of the memory stimulus by moving a cursor to the remembered location. In half the trials, randomly distributed, rTMS was applied to DLPFC, and in a separate session, to the SPL, a brain area implicated in spatialworking memorystorage. A 10Hz (3 sec, 110% of motor threshold) train of rTMS was delivered at the onset of the response period. They found that only rTMS applied to DLPFC significantly affected performance. When 10 Hz rTMS was directed to right DLPFC, accuracy declined on delayed recall trials. When 10 Hz rTMS was directed to left DLPFC, accuracy on the delayed recognition trials was enhanced. These findings support DLPFC role in memory guided response and suggest that the nature of this role varies depending on the processes required for making a response. Koch et al. used an rTMS approach to investigate the temporal dynamics and reciprocal interactions of different regions of the parietofrontal network in normal subjects performing a spatial working memory task [81]. They aimed to compare neural activity in these regions in the delay and decision phases of the task. Trains of rTMS at 25 Hz were delivered over PPC, premotor cortex (SFG), and DLPFC of the right hemisphere alternatively during the two phases. They observed TMS interference during the delay phase for PPC and DLPFC sites of stimulation, with no effect on the control site. When rTMS trains were applied during the retrieval phase, task interference was limited to DLPFC. The authors concluded their study supporting the

existence of a parallel processing network that was active during the delay phase of the task. They concluded that in DLPFC, two task specific networks coexisted: a local neural network subserving the decisional processes, and a second neural population functionally connected to PPC that was activated when spatial specific information maintained in memory was made available for use.

Luber et al. designed a delayed match-to-sample task in which rTMS at 1, 5, or 20 Hz was applied to either left DLPFC or midline parietal cortex during the delay phase of the task [94]. Only 5 Hz stimulation to the parietal site resulted in a significant decrease in reaction time with no effect on accuracy. In a second experiment, 5 Hz rTMS was directed to the parietal site during either the maintenance or retrieval phase of the recognition probe. Reaction time enhancement occurred in the retention phase but not the probe phase. They concluded their results suggest that 5 Hz rTMS may improve working memory performance when stimulation of a specific frequency was timed to a specific phase of task performance.

Earlier work to show that TMS may enhance prefrontal oscillatory synchrony in spectral bands subserving working memory information processing [7], as noted above, lends support to the idea of developing a neural circuit based model for the treatment of memory impairment using noninvasive brain stimulation (NIBS) methods such as TMS. Presently, evidence to support an NIBS based approach is supported by work to show improved clinical status in depression [110], hallucinations in schizophrenia [4, 71], and posttraumatic stress disorder [65, 113].

Visualization of the data in rTMS studies is often problematic as it requires the fusion of many data types. Fortunately, systems such as VisTrails not only provide the infrastructure for this task, but also include tools for collaboration and analysis. Unfortunately, generating and modifying these workflows requires a substantial amount of expertise. To assist scientists better manipulate the analyses and visualization, VisTrails may be used through workflow medleys [131] and VisMashups [132]. Figure 4.11 shows the VisMashup application allowing scientists to more easily interact with complex analysis and visualizations for the assessment of this rTMS experiment.

Section 4.5 has presented a method enabling scientists to visualize a large number of time frequency planes derived from EEG sensor traces, as seen in Figure 4.12. In particular, these data are used to study the dynamics of the neuronal assemblies subserving working memory in humans and how these rhythms are manipulated by repetitive transcranial magnetic stimulation. Our proposed visualization technique was used to assist in the analysis of the collected EEG data. Figure 4.13 shows a volume of the alpha frequencies generated from a single trial of the administered Sternberg test as discussed in Section 4.5.1. The high energy densities present in the volume, shown in red, indicate the activation of working memory. By ordering the volume with our greedy heuristic, the progression of activity in this band through the EEG sensor network is elucidated. In particular, the
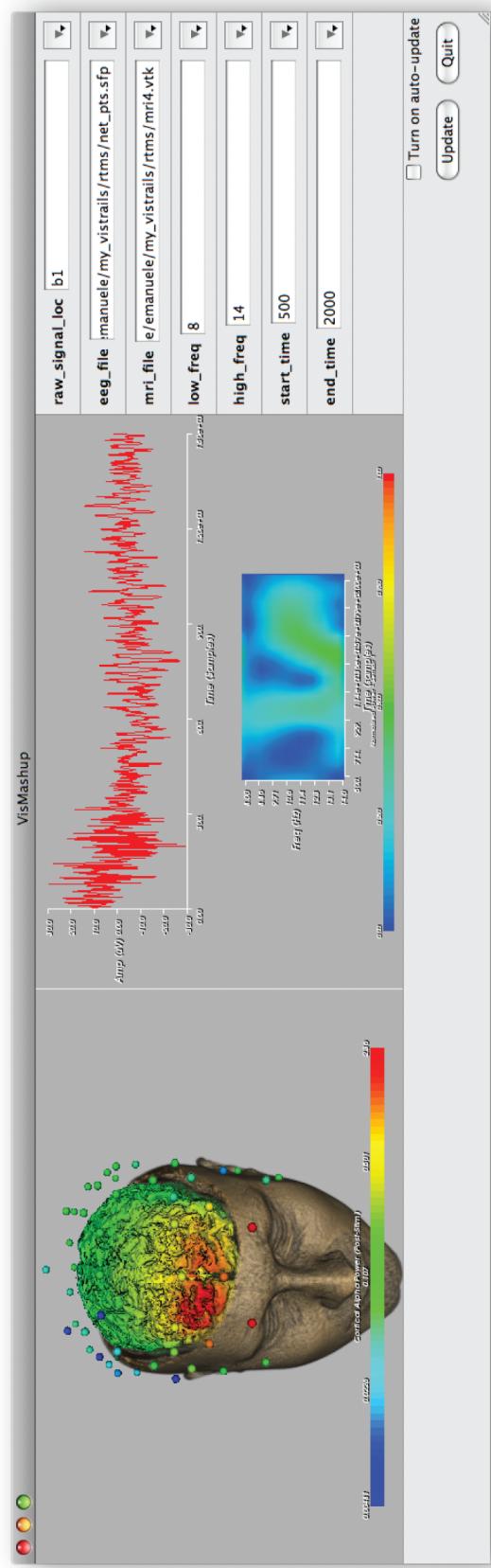
Figure 4.11: Neuroscience VisMashup. This mashup combines two pipelines—one that produces a plot for an EEG and another that creates a volume rendered visualization. To ensure that the plot and the visualization are derived for the same patient, the variables in the pipeline views corresponding to the patient (i.e., the input data set) are synchronized.

(a)   (b)
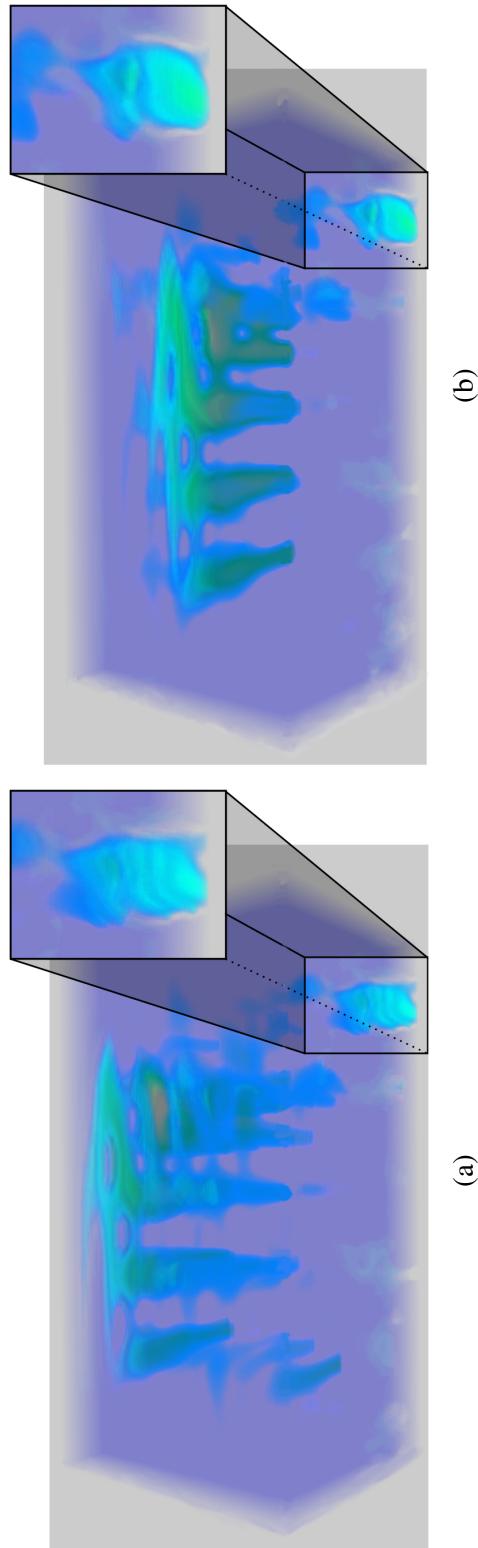
Figure 4.12: Ordering a collection of spectral planes forms a volume in which interesting features are clearly visible (highlights). Specifying a key plane that describes the features being sought (zoomed area), and examining correlations between slices, two different orderings are induced: (a) A volume using globally sorted correlations with a key plane image. (b) A volume using pair wise correlations. The volume in (b) maintains coherence across features even as planes become farther away from the user specified key plane. Without adequate coherence between planes, large features acquired at a small number of sensors appear throughout the volume as seen in (a).
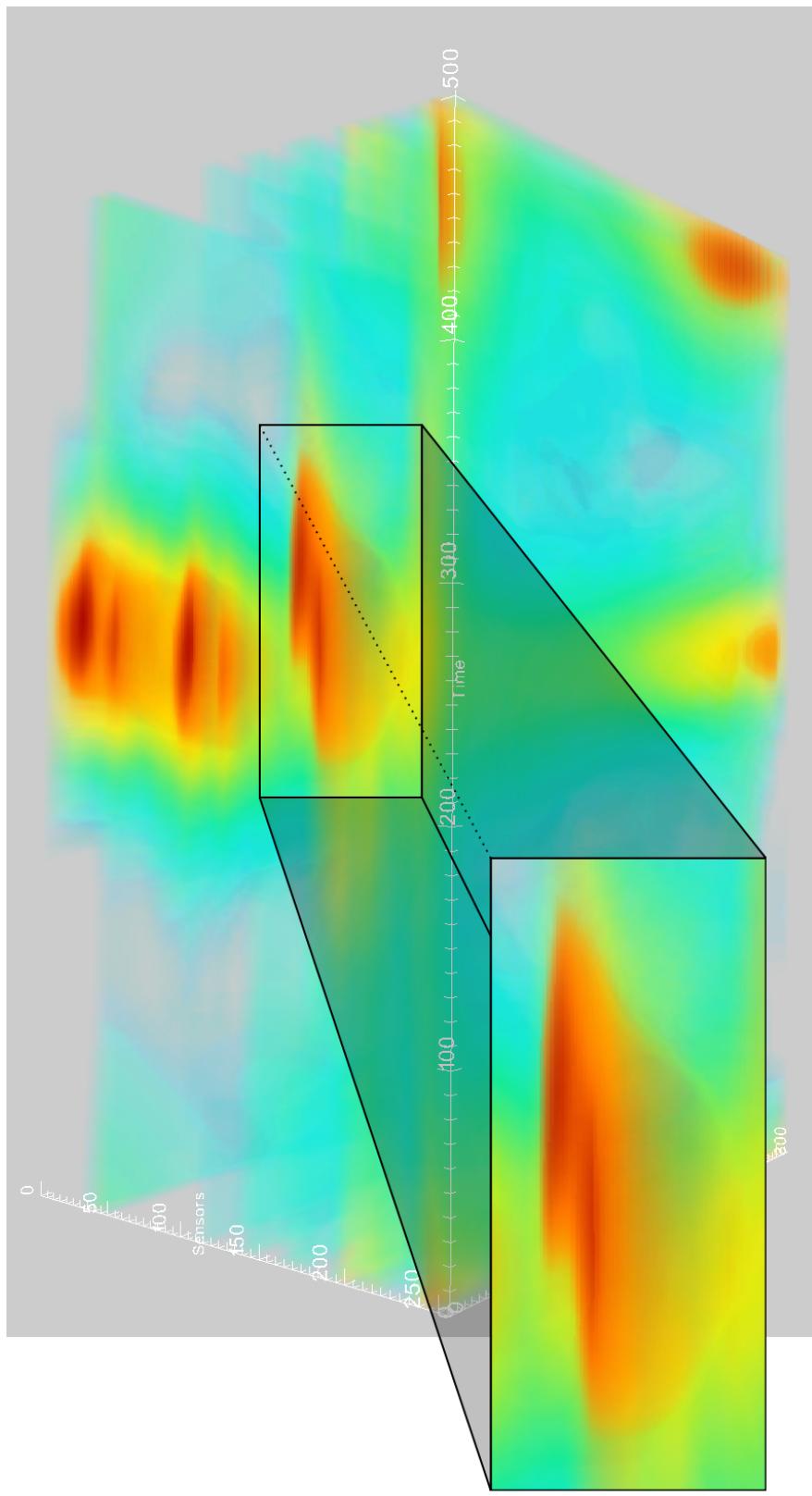
Figure 4.13: This volume represents power in the alpha frequencies for every EEG trace during a working memory task. Due to the higher coherence between features, higher power levels (colored red) can be seen moving through both time and sensor.

neural circuit is indicated by energy propagation through the volume, corresponding to the activation of different sensors.

While correlation between sensors is important to scientists studying the time frequency evolution of EEG signals, coherent representations of these planes provide additional information. Figure 4.13 depicts the movement of brain activation from sensor to sensor through time. The movement of activation is represented in the pair wise correlated volume by shifts of activation energy in the time axis. Activation patterns of this type highlight both known and potential neural pathways, or circuits, responsible for collectively processing and responding to stimuli. In the case of working memory, the highlight neural circuit shows the connectivity between the dorsolateral prefrontal, temporal and the parietal cortices predicted by Constantinidis et al. [38]. In this manner, we are able to track activation in the alpha band throughout a known circuit. Our visualization offers information regarding a range of frequencies. We exploit this property to inspect the resulting volume not just for the spectral properties of the alpha, but any time frequency patterns correlated with these activations. This is possible because they are grouped together in the volume, making them more noticeable. This correlation is important when investigating the principal effects of rTMS on the spectral dynamics of multiple frequency bands.

The visualization method for EEG data ensembles discussed in Section 4.5 is implemented as a small collection of VisTrails Modules. Because of their modular nature, this visualization technique may be added to any existing visualization using EEG data formed as a VisTrails workflow. This extension further enhances the flexibility of VisTrails as a complete EEG analysis and visualization platform.

CHAPTER 5

EVALUATING VISUALIZATION
EFFECTIVENESS USING EEG

This chapter details the method by which cognitive load imposed by visualization is measured using electroencephalography. Previous methods of measuring cognitive load during a user's task have been developed that include an additional task to perform in order to strain the cognitive system of an individual [70] so that it robustly affects the user's time to complete the operation. This method utilizes the direct inspection of brain activity to gauge the difficulty associated with interpreting different visualizations of similar data. A user study was performed to validate this approach, with all analysis steps performed in the VisTrails environment.

This chapter is organized as follows: Section 5.1 presents an overview of the techniques used to capture working memory performance and cognitive load via EEG. Section 5.2 then outlines a user study performed, with Sections 5.3 and 5.4 providing details on the extraction of cognitive load and the results of the user study, respectively. Section 5.5 concludes with a discussion of the study and results.

5.1   Working Memory

Working memory is responsible for the retrieval, manipulation, and processing of task related information and has functional importance to a variety of cognitive activities including learning, reasoning, and comprehension [14]. It is often useful to think of the working memory system in terms of a computer architecture in which working memory acts as the central processing unit (CPU) with direct connections to temporary data buffers (RAM) in the form of short term memory, and external communications (IO) through sensory perceptions and resulting reactions [14]. Of course, the actual working memory system is much more complex than a computer, and therefore dividing up the processes of the system is not always possible, as many of the functions occur across the same neural substrate [37]. Although a strict spatial segmentation of the brain in terms of working memory activity is impossible, Braver, et al. show that the working memory processing is measurable in the prefrontal cortex of the brain [25] while Constantinidis, et al. explore a more complete neural circuit for spatial working memory [38]. Working memory is also divided into

visuo-spatial, phonological, and executive subsystems [13]. In this work, our processing techniques focus on the visuo-spatial and executive working memory circuits by weighting contributions from the prefrontal cortices more heavily than those of the parietal regions.

5.1.1   Cognitive Load Theory

Cognitive load theory [148] describes the relationship between the capacity of working memory and the cognitive demands of a particular task. The core of the theory is that people have a limited cognitive capacity during learning and problem solving tasks. The way in which information is presented can affect the amount of load placed on the working memory system and thus affect performance [47]. Cognitive load theory distinguishes three types of cognitive load: *g*ermane, intrinsic, and extraneous [33]; each distinctly affecting learning and decision making. The combination of the three types characterizes the overall cognitive load [140] (Figure 5.1).

5.1.1.1   Germane cognitive load

Germane cognitive load is the load devoted to learning new cognitive schema [148]. These schema are internal representations formed in the learning process which are used over and over and may be relevant to many tasks. Once these cognitive schema are in place, the contribution of germane cognitive load to the overall load is minimal.

5.1.1.2   Intrinsic cognitive load

Intrinsic cognitive load describes the demands on working memory capacity generated by the innate complexity of the information being examined [148]. This load represents the portion of overall cognitive load that is influenced by the difficulty of the underlying task at hand and cannot be manipulated by the design of the task. An example of intrinsic cognitive load is the inherent challenge involved in adding two numbers compared to the greater challenge in solving more advanced arithmetic problems.

5.1.1.3   Extraneous cognitive load

Extraneous cognitive load measures the additional load placed on users by the design of a task [114]. This type of load can be controlled by the way information is presented [140]. For example, Figure 5.2 shows two ways to describe data. On the left is a numerical description and on the right is a visual one. The box plot quickly gives a summary of the data through a visual presentation, while the numerical display requires more extraneous cognitive load to extract the properties of the data.
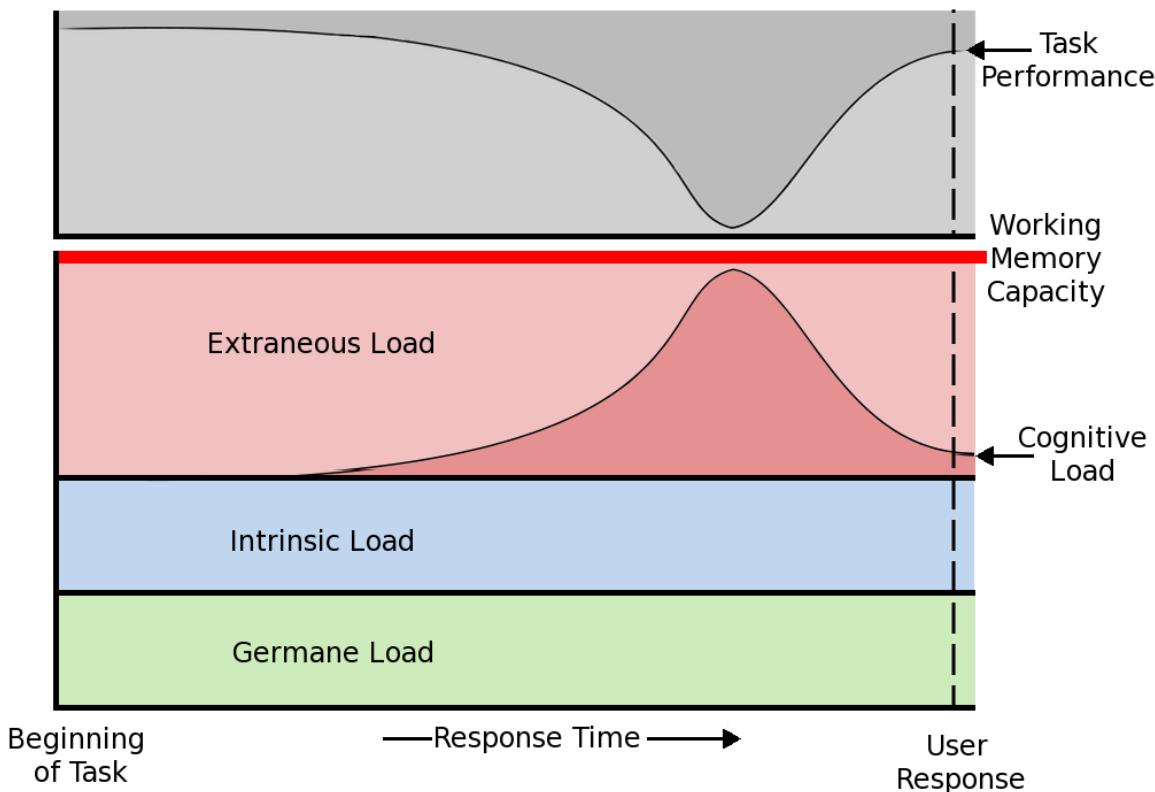
Figure 5.1: The combination of germane, intrinsic, and extraneous load to form working memory capacity and the impact of higher cognitive load (bottom curve) on task performance(top curve). Note that cognitive load peaks prior to the user's response to the task.

3, 5, 28, 78, 72, 40, 52, 37
76, 6, 26, 68, 96, 70, 66, 75
34, 33, 20, 74, 36, 85, 99, 51
99, 33, 18, 38, 14, 18, 37, 53
25, 8, 69, 85, 25, 65, 30, 28
12, 87, 59, 54, 6, 30, 16, 59
97, 66, 23, 84, 87, 76, 36, 15
97, 87, 93, 12, 70, 56, 94, 97

Table of Data   **Box Plot**
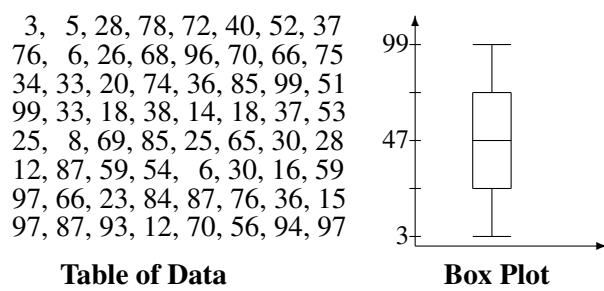
Figure 5.2: An example of extraneous cognitive load. Both figures represent the underlying data; however, the visual nature the box plot facilitates understanding by taxing the working memory system less than the numerical description.

5.1.2   Measuring Cognitive Load

One method of measuring the various types of cognitive loads is by using task completion time and accuracy. Another method of measuring cognitive load is the NASA-TLX test [70]. This test describes cognitive load in terms of subjective responses to a post experiment survey. However, EEG based processing is capable of determining cognitive load magnitude by analyzing the temporal, spectral, and spatial patterns of brain activity. The Aegis simulation environment [18] was evaluated using EEG to monitor the amplitude of brain activity induced by situational properties of the task. In this way, cognitive strains placed on the participants involved in the study were measured.

In our study, we employ EEG to measure brain activity related to cognitive load and working memory; however, other physiological measures, such as pupil dilation or galvanic skin response have also proven useful in assessing cognitive load [80, 141]. Physiological measures in user studies do not always attempt to measure cognitive stresses directly. Recently, eye tracking technology has shown great utility in studying topics ranging from graph comprehension [31] to the use of contextual cues in visualization [115]. However, it is still unclear to what degree these techniques capture cognitive responses elicited by visualization.

We exploit the spatial, temporal, and spectral organization of the neural circuits subserving working memory to measure its performance, as in [37]. The neural circuitry is monitored throughout the experiment using EEG. Although brain activity used to measure cognition is not visible in the raw EEG data, each data channel is processed to extract the spectral components associated with cognition and specifically, working memory [78]. By measuring the performance of working memory, we measure the overall cognitive load imposed on a user in realtime. This realtime measurement cannot easily distinguish one cognitive load subtype from another; however, the processing techniques allow us to make temporally sensitive analyses.

5.2   User Study of Cognitive Load

This user study is designed to evaluate different visualization techniques by measuring the amount of extraneous cognitive load each rendering imposes on the viewer. In this study, the general model of cognitive load described by Figure 5.3 is used. Because extraneous cognitive load is influenced by the way in which information is presented to the viewer, measuring its differences between visualization types provides insight into how the presentation of the data affects working memory and cognition. In order to reduce the complexity of this task, we have chosen to use simple visualization methods in this study. To this end, we compare variations of the *box plot* to see which is most effective in displaying a statistical data distribution.
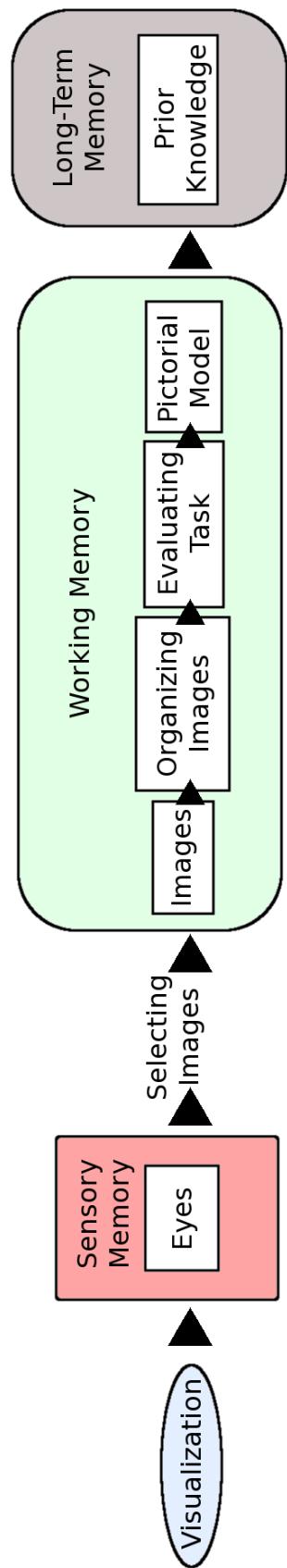
Figure 5.3: The cognitive and memory model of a single trial

The box plot is a graphical data analysis construct used to visually describe the distribution of a data set by indicating the minimum, median, and maximum data values, as well as the interquartile range (the range between the 25th and 75th percentile). The canonical box plot [155], (Figure 5.4a), does this by encompassing the central 50% of the data with a box, indicating the median with a crossbar, and extending lines out to the minimum and maximum values. Due to the box plot's simplistic representation of the underlying data, its use has become prolific in the scientific community, most notably to express error or variability within a data set. The extensive use of the box plot has supported various visual modifications, such as reducing the number of lines used to depict the plot [118, 154] (Figure 5.4b-c), or adding information about the density of the underlying data distribution [16, 69, 118] (Figure 5.4d-f).

The collection of box plots shown in Figure 5.4 were compared in this study to determine the extraneous cognitive load of each plot type. The plots were created based on 500 different normal distributions of size 100. For each distribution, the mean and standard deviation were picked uniformly random from [0,1] and [0.25, 0.75] respectively. For a single trial, two distributions are chosen and displayed using two types of box plots and the participant is asked to choose which of the distributions has a larger interquartile range.

5.2.1   Extracting Extraneous Cognitive load

EEG measures of cognition account only for overall load through the tracking of working memory performance; however, our interest lies in measuring extraneous cognitive load. In order to extract extraneous cognitive load from overall cognitive load, the design of the user study must effectively control for the other cognitive load subtypes.

Germane cognitive load is controlled for by collecting subjective data relating to participant expertise. In a post experiment survey, each participant rates their ability in interpreting the visualizations, and this information is used to approximate germane load on a per user basis. The responses to each question on the survey are given on a Likert scale [90], which asks respondents to specify their level of agreement to a statement. The survey questions are specifically designed to capture both user expertise in the interpretation of statistical data as well as the aesthetic qualities of each visualization technique. To negate the cognitive contribution of germane load, participants were required to be familiar with one-dimensional distribution data, and thus had preformed cognitive schemas. Germane cognitive load per participant was then judged to be negligible.

Intrinsic cognitive load is represented by *task difficulty*. When comparing various types of box plots, task difficulty refers to the complexity intrinsically present in deciphering differences in the interquartile range of two data sets, independent of the plotting method. When comparing images, the task is facilitated by examining common reference points within the two images. In the case
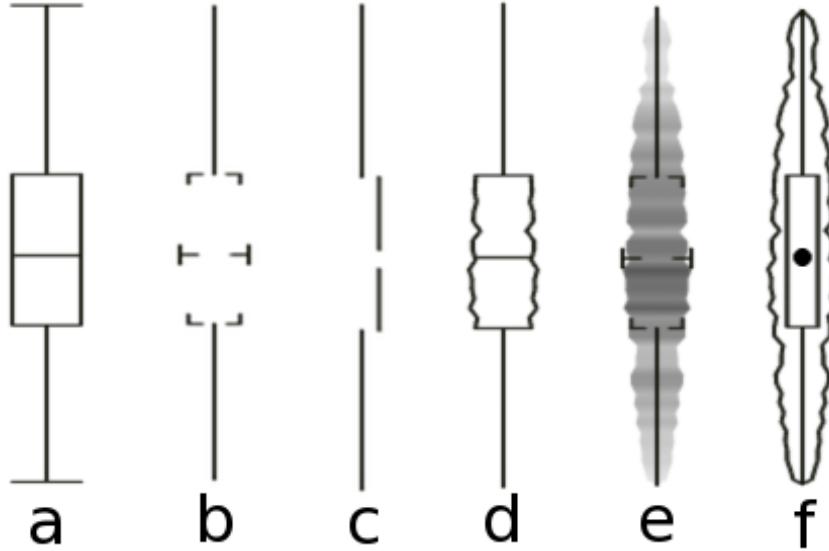
Figure 5.4: The plots used in the study. The left 3 plots are variations of the box plot: a) The Box Plot [155], b) Abbreviated Box Plot [118], c) Interquartile Plot [154]. The right 3 are box plots with additional density information: d) Vase Plot [16], e) Density Plot [118], f) Violin Plot [69].

of assessing which of two box plots has a larger interquartile range, the relevant common reference points are the locations of the first and third quartiles, and the median. The greater the similarity between the medians, the better the correspondence between the images, making the underlying task easier. However, as the distributions' interquartile ranges become similar, determining the distribution with the larger range is more difficult.

The measure of task difficulty takes into account both the interquartile range, IQR, defined as the difference between the first and third quartiles, $IQR = Q_3 - Q_1$, and the median, $\tilde{m}$, of the two underlying data distributions. Since we restrict the range of the generated distribution to be $[0,1]$, we can define task difficulty between two distributions, i and j, as $d(i,j) = 0.5(1 - |IQR_i - IQR_j| + |\tilde{m}_i - \tilde{m}_j|)$. By formulating task difficulty in this way, we are guaranteed that each single trial has a difficulty in the range $[0,1]$ in which 1 represents the highest degree of difficulty. In practice, task difficulty and thus intrinsic cognitive load was uniformly distributed in the range $[0.4, 0.8]$.

5.3   Data Analysis

Investigating the effects of different visualization techniques in terms of cognitive load requires the analysis of the various data products generated during the experiment. Time series data collected by EEG hardware must be rigorously processed to extract relevant working memory and cognitive load measures. Similarly, specific values acquired from user interaction must be manipulated to determine the task difficulty and reaction times experienced during each trial. Finally, each of

the various data products must be statistically analyzed to ensure cognitive load measures are appropriate for visualization evaluation.

5.3.1   Data Acquisition

A group of 17 individuals consisting of 10 males and 7 females participated in the user study. The user study consists of 100 independent single trials preceded by a resting period of 1-minute during which baseline values for EEG are collected. Figure 5.5 shows a participant during a single trial of the experiment. Each trial begins with a 2-second period in which no images are shown, and is followed by the display of two box plots, side by side, as shown as the stimulus at the top of Figure 5.6. The participant is asked to choose the plot with the largest interquartile range as quickly as possible, and respond by pressing the appropriate directional arrow button on a standard keyboard.

Timing and response data are recorded during the experiment through custom written display and acquisition software. A timer with 10 microsecond resolution was used to record response times during each of the single trials. In addition to the timing data used to determine reaction time, each distribution's central moments and the response given by the participant are recorded for later analysis.

Figure 5.5: A participant is fit with the EEG headset to monitor brain activity for the duration of the 100 trial experiment. Distribution visualization pairs are presented side-by-side during each trial and a keyboard is used to enter responses.
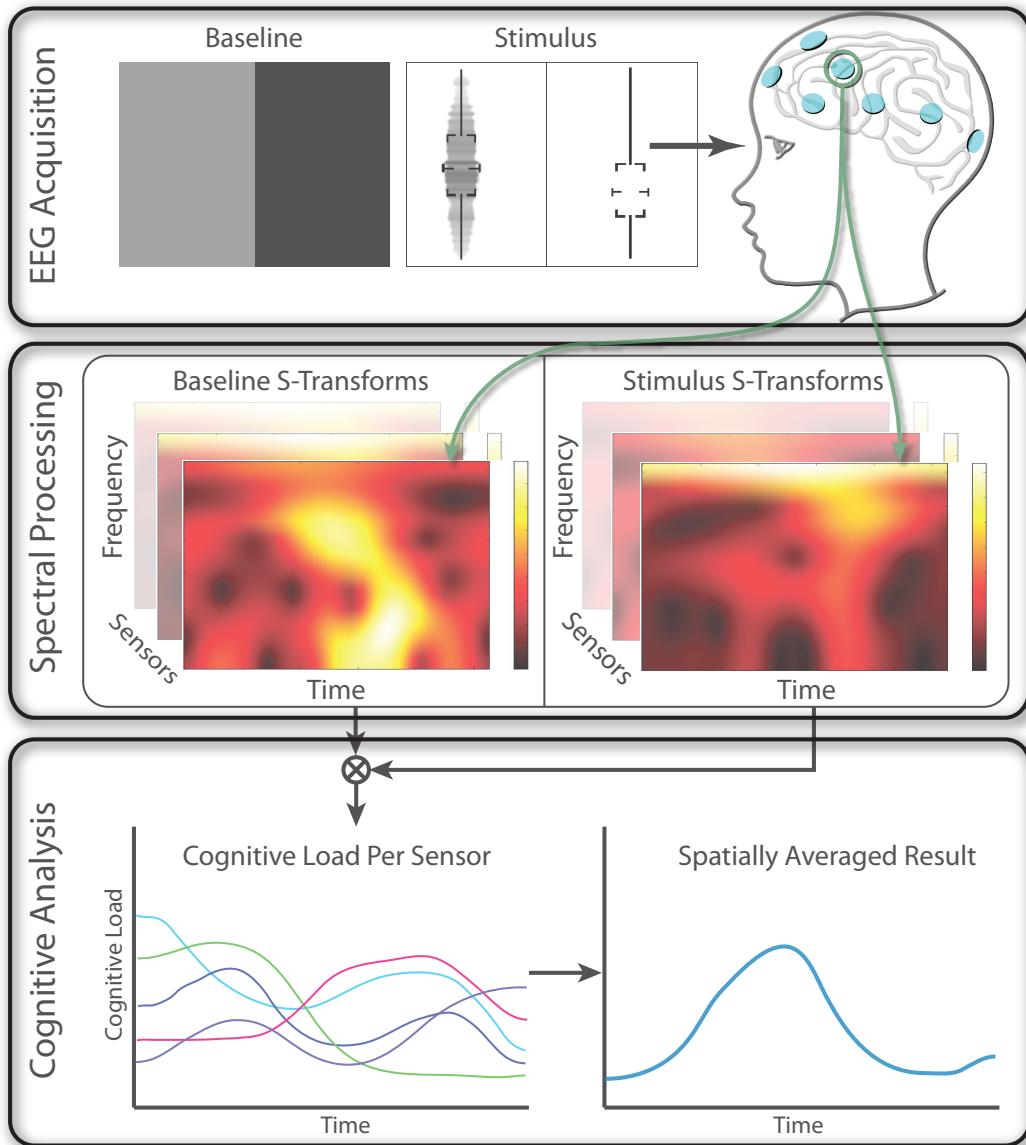
Figure 5.6: The experimental data collection and analysis workflow. EEG is collected during each of the 100 trials and then segmented into Baseline and Stimulus Epochs. These epochs are then processed using the S-Transform for each sensor. The resulting time frequency planes are further processed to extract the gravity frequency and energy density for the theta and alpha bands of frequencies in each epoch. These values are combined in the Cognitive Analysis resulting in a single time series of cognitive load for each sensor. These time series are then combined through spatially aware averaging to form the overall cognitive load for the trial.
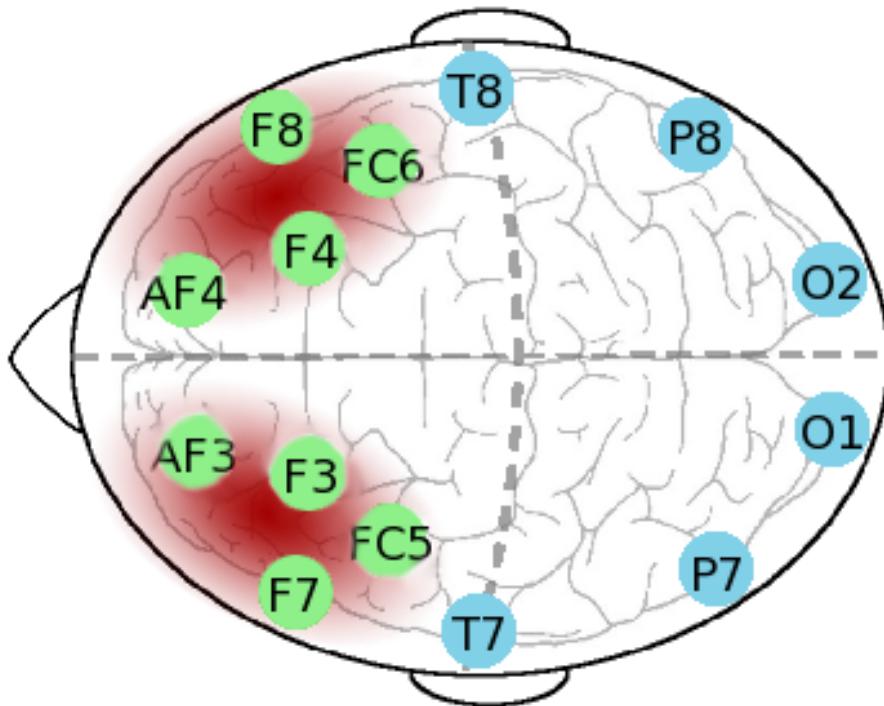
Figure 5.7: Sensor placement around the prefrontal cortex of the 14 data channels in the Emotiv EEG. The regions in red show the Gaussian weighting used to emphasize the regions of the brain most related to working memory.

EEG data are collected at 128 Hz from an Emotiv EPOC wireless EEG headset (`http://www.emotiv.com`). The Emotiv headset exposes 14 data channels with two bipolar reference electrodes spatially organized using the International 10–20 system, as seen in Figure 5.7. The Emotiv Software Development Kit (SDK) provides a packet count functionality to ensure no data are lost, a writable marker trace to ease single trial segmentation tasks, and realtime sensor contact to ensure quality measurements.

During the experiment, a unique marker value is inserted into the marker trace to signal the end of the 1-minute resting period. Additional markers are inserted each new trial, the presentation of each pair of distributions, and the user response which signals the end of a single trial. The EEG record is then segmented, using the marker trace, into the resting segment, used as a baseline measurement of brain activity, and 100 single trials. A single trial includes a 2-second resting period used to form intertrial baseline measurements, followed by the presentation of the distribution pair. Each trial may be of variable length due to reaction time differences, so a window of 1.0 seconds surrounding the user response forms the trial epoch.during the study.

Since our cognitive load measure is computed from EEG, care must be taken to account for the spatial organization of the brain. Rowe, et al. discuss the roll of the prefrontal cortex of the brain in various aspects of working memory [125]. The spatial activation sites were found to be quite localized; however, EEG experiences volume conduction causing activity generated at a single point to be measured at multiple sensors. To help account for this, spatial averaging was performed using Gaussian weights centered at the prefrontal cortex on each brain hemisphere defined in the 10-20 electrode placement system, as shown in Figure 5.7. The parametrization of the Gaussian was set to encompass sensors F7 and F3 and their contralateral pair F4 and F8 in the first standard deviation. There were no substantial differences between the left and right hemispheres found during later analysis.

5.3.2 EEG Signal Analysis

The first step in processing the raw EEG signals is to segment the 14 time series (one for each sensor) into individual trials. Next, each trial is divided into the intertrial baseline and the trial stimulation. Both of these tasks use the markers inserted into the EEG record, as discussed in Section 5.3.1. The baseline and stimulus signals are then transformed, using the S-Transform to determine the power change and frequency shift induced by the stimulation. These values are used to calculate the cognitive load experienced at each of the 14 sensors for the trial in question. Spatially averaging these 14 values gives a single measurement for cognitive load. Figure 5.6 shows the workflow of the experiment from data collection through analysis.

5.3.2.1 Artifact Detection and Removal

Since EEG measures voltages at the scalp, there are many possible sources for data contamination that must be addressed. Artifacts related to eye blinks and other muscle movements in addition to physical movements of the sensors themselves must be removed before the EEG traces can be processed. We have adapted work by Berka, et al. to decontaminate EEG signals generated by Emotiv hardware [17] and rely on the Emotiv SDK to automatically detect eye blinks. Since muscle contraction and control are generally governed outside of the frequency range of interest [130], we are able to use frequency band limiting procedures such as low pass, high pass and notch filters to adequately remove these signal components. If, after removing EEG artifacts, the energy densities of the alpha or theta frequency bands are changed by more than 20% of their original values, the trial is removed from all further analysis. This criterion is informed by the bad channel removal method discussed by Anderson, et al. [7]. In this study, we threw out 3% of the trials due to excessive signal degradation from movement and 1.5% due to high change in spectral densities, totalling 4.47% of the total trials being removed from further analysis.

5.3.2.2 Spectral Decomposition of Cognitive Load

In order to understand cognitive load, we must examine the spectral characteristics of the EEG signals. Based on the work of Klimesch [78], we focus our analysis on the alpha (7.5 – 12.5 Hz) and theta (4 – 7.5 Hz) frequency bands, which have been identified as reflecting cognitive and memory performance. We use the S-Transform [147] to decompose the signal into an appropriate time frequency representation. The S-Transform was chosen over other transformations because it offers adaptive spectral and temporal resolution similar to the Wavelet Transform and is a direct mapping to the complex Fourier Domain.

To be able to properly assess the spectral evolution of EEG associated with working memory, each trial is processed with respect to its own intertrial rest period. The individual alpha and theta frequencies are determined for both the trial and rest period and their amplitudes measured [78]. By comparing these values, a shift of both the individual frequencies as well as their amplitudes are revealed. The degree of change in these amplitudes, weighted by the amount of shift in the frequency domain, determine the working memory and cognitive load characteristics for each single trial, as described in Equation 5.2.

Our computation of cognitive load derived from EEG uses the individual mean frequencies in both the alpha and theta frequency bands. The mean frequency is computed as:

$$f(\omega) = \frac{\sum_{i=0}^{n-1} I_{\omega(i)} f_{\omega(i)}}{\sum_{i=0}^{n-1} I_{\omega(i)}} \tag{5.1}$$

where ω is the frequency band in question, n is the number of frequency bins in ω, f_i is the frequency at bin i and I_i is the energy density of ω at frequency bin i. This formulation of mean frequency is used to compute the frequency shifts in both the alpha and theta wavebands. The frequency shift of a waveband is given by $f_t(\omega) - f_b(\omega)$ where f_t is the frequency content determined from EEG collection during each trial and f_b is the frequency content collected during intertrial rest periods. Additionally, the change in energy density in a waveband, $\Delta|f(\omega)|$, is the difference of energy densities at the mean frequencies: $\Delta|f(\omega)| = |f_t(\omega)| - |f_b(\omega)|$.

Klimesch identified working memory performance decreases during task related stimulation expressed as theta power decreases with simultaneous alpha power increases with respect to baseline measurements [78]. We form our model of cognitive load per trial, $L(t)$, as the combination of frequency and power changes in both the alpha and theta bands.

$$L(t) = \Delta|f_t(\alpha)|f_t(\alpha) - \Delta|f_t(\theta)|f_t(\theta) \tag{5.2}$$

5.4   Cognitive Load User Study Results

Using direct inspection of brain activity during a visualization task provides us with additional empirical data regarding the effectiveness of different rendering methods. Because EEG measurements are not corrupted by the participant's subjectivity or the benefit of hindsight, as may be the case during post experiment surveys, they are well suited for determining the effectiveness of visualization.

Based on our EEG recordings and subsequent analysis, the canonical Box Plot was found to place the least strain on the user's cognitive resources for the task at hand. Table 5.1 shows the computed cognitive load for each plot type using both Gaussian and constant spatial averaging. The table indicates the Box plot and the Density Plot incurred the lowest cognitive load scores (in bold) using Gaussian and constant weighting, respectively. This result highlights the effect of the spatial averaging on overall cognitive load. Using Gaussian weights helps account for the brain's natural spatial organization, providing a more reliable measure. Interestingly, the Violin and Interquartile plots induced the highest cognitive load (in italics). This may be due to greater visual complexity or the reduction of distinguishable visual elements; however, the validation of such claims warrants additional study.

Reaction time is important in determining working memory performance and capacity [146]. While reaction time cannot measure working memory performance directly, it is an appropriate means of capturing the aggregated performance and capacity of working memory. As the role of reaction time in determining working memory performance is well explored [7, 120], we focus our analysis on the assessment of brain activity via EEG processing.

Figure 5.8 plots the computed cognitive load and the reaction time from this experiment against the task difficulty for each trial spanning all participants in the user study. The figure suggests correlation between task difficulty and both reaction time as well as the measurement of cognitive load; as the difficulty of the task increases, so does the computed cognitive load and reaction times. However, there is a relatively large variance in both cognitive load and reaction times, particularly in the investigation of high difficulty tasks. One explanation for this large variance is an incorrect model for task difficulty. The computed task difficulty (Section 5.3) uses only the median and interquartile range of each distribution. Exploring different formulations for task difficulty may result in a more robust correlation between each trial's computed difficulty and the cognitive load computed. Additionally, our cognitive load measure weights contributions from the alpha and theta frequencies equally. It is possible that a more advantageous combination of theta and alpha spectral changes exists, but adequately exploring the nuances of these formulations is beyond the scope of this paper.
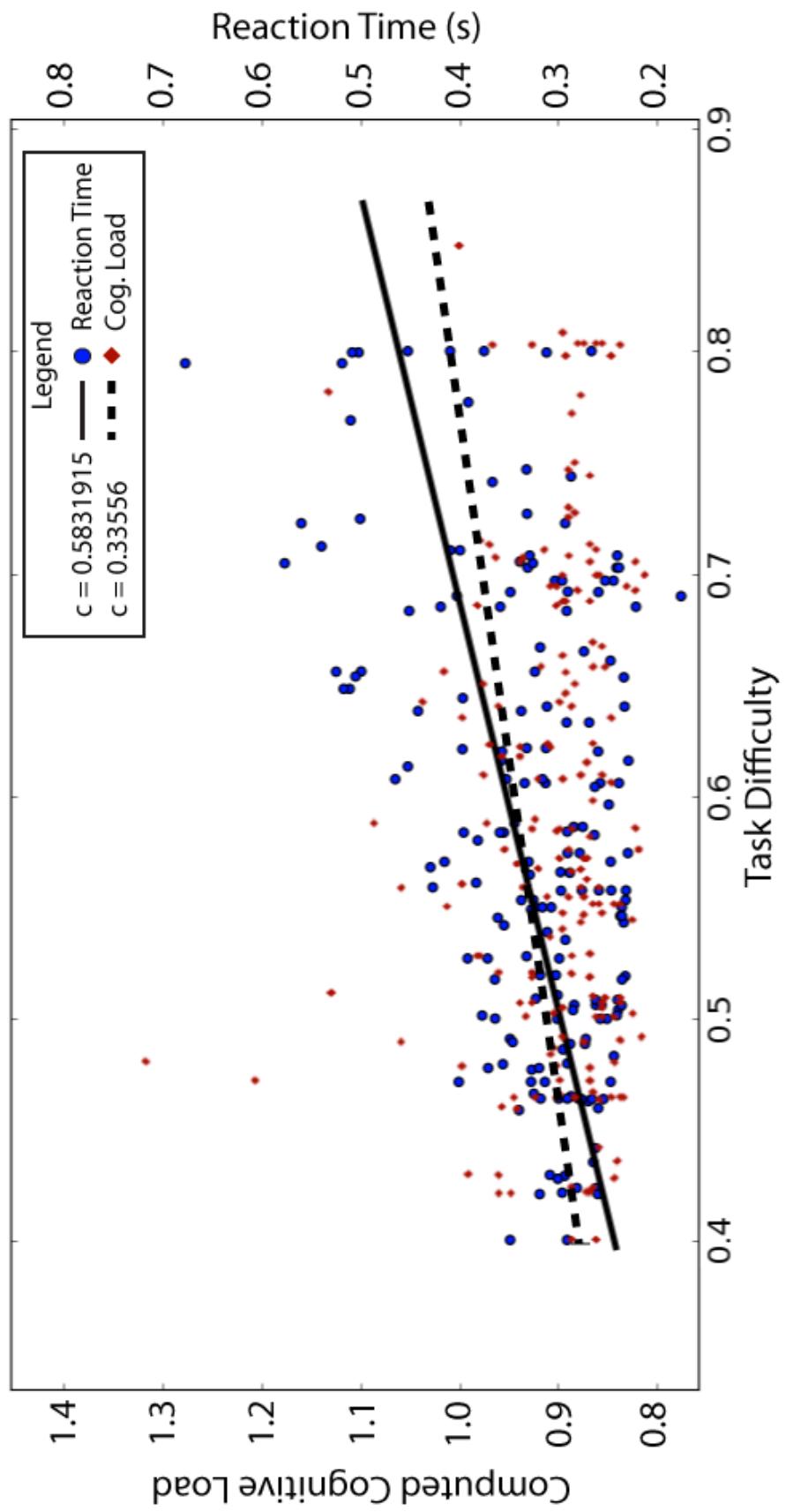
Figure 5.8: Results from this experiment suggest a correlation between greater task difficulty and higher cognitive load. Here, task difficulty is plotted against computed cognitive load and reaction time for each valid trial across all participants.

Table 5.1: Computed cognitive load for each plot type. Constant and Gaussian spatial averaging are shown. Lowest cognitive load scores are highlighted in bold while highest scores are italicized.

	Box	Abbrv.	Interquartile	Vase	Density	Violin
Constant	1.101	1.284	1.214	1.571	**0.830**	*1.619*
Gaussian	**0.815**	0.833	*1.563*	1.203	1.285	1.492

5.4.1   Statistical Analysis

In order to determine significant correlation between the measured data and visualization type, we employ paired 2 tailed T-tests. T-tests were used to determine significance of spectral properties departing from baseline measurements taken as well as spectral differences between visualization types. All statistical tests used the null hypothesis that there is no significant change between the two distributions being analyzed. Each distribution tested was inspected to verify it was not multimodal prior to analysis.

Table 5.2 displays the maximum significance values (p values) as computed for cognitive load by the 2 tailed T-tests. Of particular interest are the high degrees of similarity between the Box Plot and Abbreviated Box Plot (Box and Abbrv. in Table 5.2) and the Violin and Interquartile Plots (Violin and Interquartile in Table 5.2). All tests were performed with cognitive loads computed using Gaussian weights as discussed in Section 5.3.1.

5.5   Discussion

In this study, we explored different methods of visualizing distribution data. For each method under consideration, the cognitive load associated with interpreting the interquartile range was determined. While each of the visualizations used for this study displayed the interquartile range of a distribution in some way, not each rendering displayed the same amount of data associated with

Table 5.2: Pairwise significance values for cognitive load of the Box Plot (Box), Abbreviated Box Plot (Abbrv.), Interquartile Plot (Interquartile), Vase Plot (Vase), Density Plot (Density) and Violin Plot (Violin). While most significance values are below 0.01, some pairs of comparisons generated similar distributions. The Box Plot and abbreviated version score similarly as do the Interquartile and Violin Plots.

	Box	Abbrv.	Interquartile	Vase	Density
Violin	0.001	0.001	0.134	0.0015	0.0015
Density	0.001	0.001	0.003	0.002	x
Vase	0.001	0.001	0.0015	x	x
Interquartile	0.001	0.001	x	x	x
Abbrv.	0.216	x	x	x	x

each underlying data set. For example, the Violin Plot rendered the sample density as described by its histogram whereas the Box Plot did not. These differences enable a different set of questions to be asked about these visualizations that cannot be asked about other visual representations. This study, like others, focuses on the effectiveness of visualization method with respect to a single subset of appropriate interpretation tasks.

Until recently, the expense of EEG technology greatly limited its application in the field of user studies. The Emotiv EPOC headset used in this experiment provided a cost effective means of EEG acquisition. However, although this system conforms to the international 10–20 standard for electrode placement, getting each electrode in the proper position is important and nontrivial. Additionally, the analysis and interpretation of EEG data remains difficult, requiring training and expertise.

The visualizations and interpretations required during this user study were purposefully chosen to be elementary. The simplicity of this study allowed participants to be chosen from a wide range of potential candidates in order to minimize the potential for schema creation and over representation of germane cognitive load. In addition to controlling germane cognitive load, this decision allowed us to more completely regulate and estimate the contribution of intrinsic cognitive load during each single trial. By acknowledging and controlling these two parameters, we were able to more thoroughly process the resulting data without substantially complicating the analysis.

Minimizing the visualization and task complexity eased requirements for the analysis and processing steps used in this study; however, the experimental design was still difficult. After determining the appropriate visualizations to use during the experiment, finding the proper interpretation task proved to be arduous. Using too simple an interpretation task did not create enough cognitive load to substantially influence working memory performance. Meanwhile, employing too complex a task induced cognitive overload, complicating analysis. Cognitive overload was identified by the movement of the individual alpha frequency outside of the 8–12 Hz band of frequencies, following the results of Klimesch and Gevins, et al. [57, 78].

Much work has been done to explore the effects of practice on cognitive measures, as the introduction of these effects often confound analysis. Berry, et al. [19] found that practice does not expand the capacity of working memory and cognition, as was previously thought, but instead improves the efficiency of data encoding. This finding implies that the inverse relationship between available working memory capacity and cognitive load is maintained regardless of practice during an experiment. The spectral dynamics of practice effects in cognition were explored by Gevins, et al. [58]. Practice was found to decrease reaction time, but also increase spectral organization. The spectral changes induced by practice comprised an increase in power and frequency modulation

prior to the task onset. To mitigate the effects of practice in this study, we reevaluate baseline conditions during the rest period before each trial begins. While this helps minimize the practice effect in analysis for this study, reevaluating baseline performance may not be possible in more complex, or time sensitive experiments.

The temporal, spatial, and spectral organization of brain activity enable both analysis and interpretation. Despite an adequate tool set for the processing and general analysis of EEG signals, their interpretation requires domain experts. The multidisciplinary nature of this study was essential for proper examination of the results we collected. Without the close collaboration between computer scientists, neuroscientists, and psychiatrists, the success of this study would have been jeopardized.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

The use of the provenance management system, VisTrails, for the analysis and visualization of EEG data from two studies of working memory have been presented in this dissertation. Although VisTrails is not the only system used for EEG data analysis [41], it is currently the only one that was designed to capture full provenance of any modification made to the underlying processing workflow. Chapter 3 details the use of EEG analysis within VisTrails, including using external, opimized processing libraries. These techniques were then applied to data collected during a study of working memory in which repetitive transcranial magnetic stimulation (rTMS) was used to manipulate the neural circuits subserving the working memory process discussed in Chapter 4. It was found that working memory was robustly manipulated by rTMS in a way measurable by EEG. This induced change of working memory performance may be leveraged as part of a treatment regimen for people suffering from debilitating neurological disorders such as schizophrenia [7].

Another study was then performed to examine the links between working memory performance, cognitive load, and the difficulty inherent to interpreting visualizations in Chapter 5. In this study, VisTrails-provided functionality was the only system in which data were analyzed. This study found that data collected from EEG may be processed to form a measure of cognitive load associated to the visual representation of data. By comparing the cognitive loads induced by various types of visualizations of similar data, a measure of visualization efficacy is formed. However, this efficacy score is valid only for the specific set of questions posed during the experiment.

There are several avenues for future work that are immediately apparent. As the VisTrails environment is continually updated, new functionality and provenance enabled operations become possible. Evaluating the utility of these functions with respect to EEG and other biomedical data must be done to properly incorporate functionality into existing collaborations. Furthermore, advances in signal processing, visualization, and the understanding of EEG interpretation will necessarily require further extension of the module packages developed for VisTrails. Without these updates, VisTrails as an EEG processing and visualization platform will stagnate.

The consequences of the findings discussed in Chapter 4 may be profound, but require more work to properly verify. The application of rTMS to healthy individuals was shown to have robust

positive effects on working memory performance. Individuals suffering from neuropsychiatric disorders containing underperforming or malfunctioning working memory may benefit from rTMS as a treatment. Although pharmaceutical methods exist for the treatment of individuals with diseases such as schizophrenia, these options neither address the root causes of the disease nor guarantee results. Treatment of these neurological disorders via rTMS may provide a better alternative to patients who are unresponsive to pharmacological treatments or may increase the efficacy of the drugs used in them. However, as this method as a potential treatment is in its infancy, several more studies in this direction are necessary.

The assessment of visualization through the examination of EEG data is an entirely new approach to the field. The study presented in Chapter 5 used only the simplest of data and visualization techniques. Employing more visualization methods, including 2D and 3D renderings, color, and more complex datasets is necessary. The results of these studies should be stored in a database in order to then catalogue and compare the features of visualization that may increase or decrease cognitive loads.

A database of visualization effectiveness containing many different tasks and visualizations would likely provide additional insights into the nature of visualization. In particular, such a database may help to classify different visualization methods by linking features highlighted by a particular technique with the cognitive load they impose. The classification of visualizations in this way may assist researchers striving to formalize the field of visualization.

Finally, the link between working memory, cognition, and EEG may further be exploited in a variety of ways. Determining methods with lower cognitive costs associated with them may benefit several different fields. Studies using EEG to evaluate teaching techniques may lead to more effective, or more personalized, teaching tools. Studying the effects of cognitive overload may lead to the development of new and better coping mechanisms. Better coping with cognitive overload will likely lead to faster knowledge generation. While these possibilities exist for future study, they all have a common prerequisite. Each avenue for future work depends entirely on more studies using EEG and cognition to be performed.

APPENDIX

EXPERIMENTAL REPRODUCIBILITY

The use of VisTrails as an EEG processing platform was discussed in Chapter 3. Although this chapter explored the use of the Brainiac VisTrails package for signal analysis, its functionality is derived from the modules in more basic packages. The Brainiac, NumSciPy, and VisualAsses packages have been made available under the Lesser GNU General Public License (LGPL). The source code for these VisTrails packages and the VisTrails associated with the figures in this dissertation are available at:

```
http://www.sourceforget.net/projects/vt-brainiac
```

A.1   Abstractions Over VisTrails Workflows

The Brainiac VisTrails package discussed in Chapter 3 contains modules that represent abstractions over partial workflows. These abstractions allow users to more easily create complex workflows out of simpler functional components. As pointed out in Chapter 6, Brainiac does not make full use of recent additions to VisTrails, most notably the increased support for abstractions called Subworkflows.

Utilizing the subworkflow support in new versions of VisTrails simplifies the construction of abstractions useful to EEG analysis and visualization. While the existing Brainiac package uses abstraction at the code level, as seen in Listing A.1, new abstraction mechanisms support the creation of new modules using the VisTrails interface (Figure A.2). Although the Brainiac package supports abstraction versioning via package versioning, this system is both brittle and inefficient. Since the new VisTrails subworkflow system utilizes a full VisTrails representation for each abstraction, each new module maintains its own version history and may be interchanged in workflows at will.

Adopting the new VisTrails subworkflow system increases the robustness of Brainiac abstractions. As subworkflows are constructed from within the VisTrails environment, all of the benefits gained from using VisTrails become available to the Brainiac package. These include per abstraction provenance and versioning, modification of abstractions through analogy, and the rich querying mechanisms of VisTrails.

```python
class BrainiacVolumeRender(BrainiacRender, Module):
    def process_compute(self):
        if self.vol_texture:
            self.mapper = vtk.vtkVolumeTextureMapper3D()
        else:
            self.mapper = vtk.vtkVolumeRaycastMapper()

        self.mapper.SetInputConnection(
            self.get_data().vtkInstance.GetProducer().GetOutputPort())
        self.vprop = vtk.vtkVolumeProperty()
        self.tf.set_on_vtk_volume_property(self.vprop)
        self.volume = vtk.vtkVolume()
        self.volume.SetMapper(self.mapper)
        self.volume.SetProperty(self.vprop)
        if self.plane != None:
            self.mapper.AddClippingPlane(self.plane.plane.vtkInstance)
            self.plane.widget.vtkInstance.SetProp3D(self.volume)
            bounds = self.volume.GetBounds()

            c_x = (bounds[1] - bounds[0]) / 2.0
            c_y = (bounds[3] - bounds[2]) / 2.0
            c_z = (bounds[5] - bounds[4]) / 2.0
            self.plane.widget.vtkInstance.SetOrigin(c_x, c_y, c_z)

            self.plane.widget.vtkInstance.SetPlaceFactor(1.0)
            self.plane.widget.vtkInstance.PlaceWidget()

        return (self.volume,)
```

[pt]

Figure A.1: An example of Abstraction within the current framework of the Brainiac VisTrails Package. Notice that each of the concrete modules in the abstracted workflow must be created and connected within the module's source code.
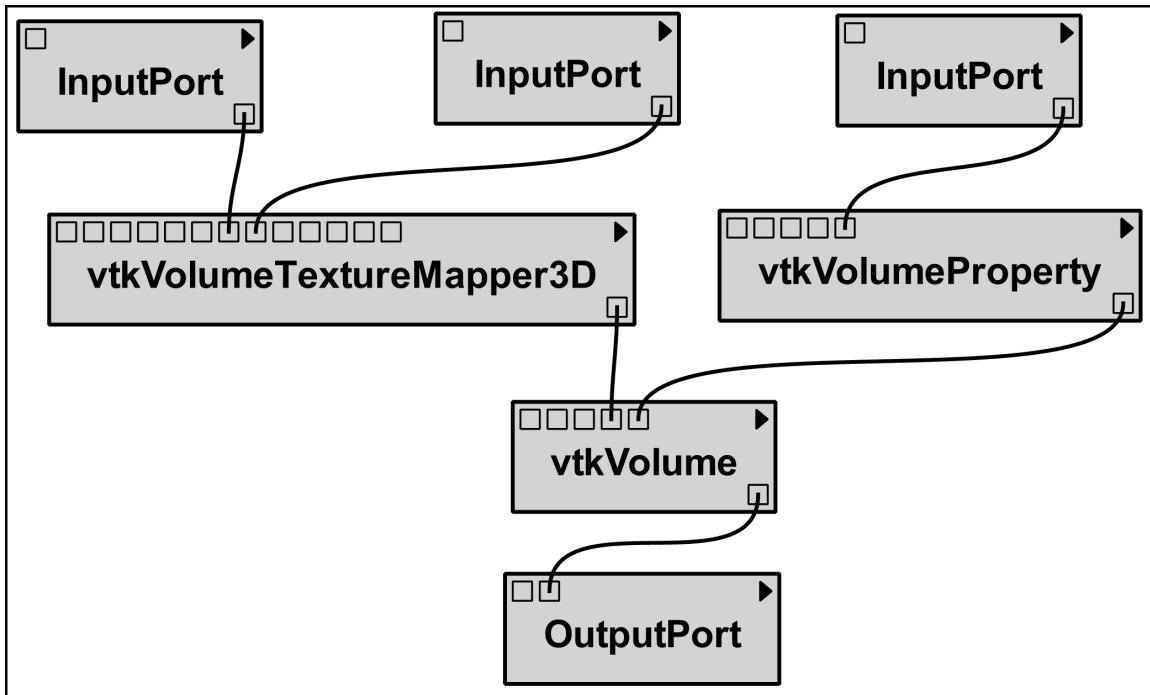
Figure A.2: An example of Abstraction using VisTrails subworkflow system. Notice that the subworkflow contains InputPort and OutputPort Modules to provide a concise interface and enhance their ease of use.

A.2 Reproducible Analysis

The reproducibility of EEG analyses is enabled by the provenance tracking of VisTrails. Most of the analyses throughout this dissertation are reproducible within VisTrails. The VisTrails files used to create the visualizations and analyses may be found at:

```
http://www.sourceforget.net/projects/vt-brainiac
```

However, due to the fact that EEG and MRI data is collected from human participants, our IRB protocol does not allow us to redistribute the source data. This makes the strict reproduction of the results contained in this document not possible.

REFERENCES

[1] The vistrails project. `http://www.vistrails.org`.

[2] AACH, J., AND CHURCH, G. M. Aligning gene expression time series with time warping algorithms. *Bioinformatics 17*, 6 (2001), 495–508.

[3] AIGNER, W., MIKSCH, S., MÜLLER, W., SCHUMANN, H., AND TOMINSKI, C. Visualizing time-oriented data – A systematic view. *Computers and Graphics 31*, 3 (2007), 401–409.

[4] ALEMAN, A., SOMMER, I. E., AND KAHN, R. S. Efficacy of slow repetitive transcranial magnetic stimulation in the treatment of resistant auditory halluncinations in schizophrenia: A meta-analysis. *Journal of Clinical Psychiatry 68* (2007), 416–421.

[5] ALLEN, J. Short term spectral analysis, synthesis, and modification by discrete Fourier Transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing] 25*, 3 (Jun 1977), 235–238.

[6] ANDERSON, E. W., POTTER, K. C., MATZEN, L. E., SHEPHERD, J. F., PRESTON, G. A., AND SILVA, C. T. A user study of visualization effectiveness using EEG and cognitive load. *Computer Graphics Forum 30*, 3 (2011).

[7] ANDERSON, E. W., PRESTON, G. A., AND SILVA, C. T. Towards development of a circuit based treatment for impaired memory: A multidisciplinary approach. *IEEE EMBS Special Topics on Neural Engineering* (2007).

[8] ANDERSON, E. W., PRESTON, G. A., AND SILVA, C. T. Using pythong for signal processing and visualization. *IEEE Computing in Science and Engineering 12*, 4 (2010).

[9] ANDERSON, E. W., SILVA, C. T., AHRENS, J. P., HEITMANN, K., AND HABIB, S. Provenance in comparative analysis: A study in cosmology. *IEEE Computing in Science and Engineering 10*, 3 (2008), 30–37.

[10] ANDREASEN, N. C., NOPOULOS, P., O'LEARY, D. S., MILLER, D. D., WASSINK, T., AND FLAUM, M. Defining the phenotype of schizophrenia: Cognitive dysmetria and its neural mechanisms. *Biological Psychiatry 46* (1999), 908–920.

[11] ANGELAKIS, E., AND STATHOPOULOU, J. Electroencephalographic peak alpha frequency correlates of cognitive traits. *Neuroscience Letters 371*, 1 (2004), 60–63.

[12] ANTONINI, M., BARLAUD, M., MATHIEU, P., AND DAUBECHIES, I. Image coding using wavelet transform. *Image Processing, IEEE Transactions on 1*, 2 (Apr 1992), 205–220.

[13] BADDELEY, A. *Working Memory*. Oxford University Press, 1983.

[14] BADDELEY, A. Working memory: The interface between memory and cognition. *J. of Cognitive Neuroscience 4*, 3 (1992), 281–288.

[15] BANQUET, J. Spectral analysis of the EEG in meditation. *Electroencephalography and Clinical Neurophysiology 35*, 2 (1973), 143–151.

[16] BENJAMINI, Y. Opening the box of a boxplot. *The American Statistician 42*, 4 (1988), 257–262.

[17] BERKA, C., LEVENDOWSKI, D. J., CVETINOVIC, M. M., PETROVIC, M. M., DAVIS, G., LUMICAO, M. N., ZIVKOVIC, V. T., POPOVIC, M. V., AND OLMSTEAD, R. Real-time analysis of EEG indices of alertness, cognition and memory with a wireless EEG headset. *Int'l J. of HCI 17* (2004), 151–170.

[18] BERKA, C., LEVENDOWSKI, D. J., RAMSEY, C. K., DAVIS, G., LUMICAO, M. N., STANNEY, K., REEVES, L., REGLI, S. H., TREMOULET, P. D., AND STIBLER, K. Evaluation of an EEG-workload model in the Aegis simulation environment. *Proceedings of SPIE* (2005), 90–99.

[19] BERRY, S., ZANTO, T. P., RUTMAN, A. M., CLAPP, W. C., AND GAZZALEY, A. Practice-related improvement in working memory is modulated by changes in processing external interference. *J. of Neurophysiology 102* (2009), 1779–1789.

[20] BOROOJERDI, B., PHIPPS, M., KOPYLEV, L., WHARTON, C. M., COHEN, L. G., AND GRAFMAN, J. Enhancing analogic reasoning with rTMS over the left prefrontal cortex. *Neurology 56* (2001), 526–528.

[21] BOSE, R., AND FREW, J. Lineage retieval for scientific data processing: A survey. *ACM Computing Surveys 37*, 1 (2005), 1–28.

[22] BOWERS, S., MCPHILLIPS, T. M., AND LUDÄSCHER, B. Provenance in collection-oriented workflows. *Concurrency and Computation: Practice and Experience 20* (2008), 519–529.

[23] BRACEWELL, R. *The Fourier Transform and Its Applications*. McGraw-Hill, New York, U.S.A., 1965.

[24] BRANDT, S. A., PLONER, C. J., MEYER, B. U., LEISTNER, S., AND VILLRINGER, A. Effects of repetitive transcranial magnetic stimulation over dorsolateral prefrontal and posterior parietal cortex on memory-guided saccades. *Experimental Brain Research 118* (1998), 197–204.

[25] BRAVER, T. S., COHEN, J. D., NYSTROM, L. E., JONIDES, J., SMITH, E. E., AND NOLL, D. C. A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage 5*, 1 (1997), 49 – 62.

[26] BRIGNANI, D., MANGANOTTI, P., ROSSINI, P. M., AND MINIUSSI, C. Modulation of cortical oscillatory activity during transcranial magnetic stimulation. *Human Brain Mapping 29* (2008), 603–612.

[27] BRYSON, G. Initial and final work performance in schizophrenia: Cognitive and symptom predictors. *Journal of Nervous and Mental Disorders 191*, 2 (2003), 87–92.

[28] BUST, J., AND GALBRAITH, G. EEG correlates of visual-motor practice in man. *Electroencephalography and Clinical Neurophysiology 38*, 4 (1975), 415–422.

[29] CALLAHAN, S. P., FREIRE, J., SANTOS, E., SCHEIDEGGER, C. E., SILVA, C. T., AND VO, H. T. Managing the evolution of dataflows with vistrails. *Data Engineering Workshops, 22nd International Conference on 0* (2006), 71.

[30] CALLICOTT, J., VERCHINSKI, B., MARENCO, S., EGAN, M., AND WEINBERGER, D. Complexity of prefrontal cortical dysfunction in schizophrenia: More than up or down. *American Journal of Psychiatry 160*, 12 (2003), 2209–2215.

[31] CARPENTER, P. A., AND SHAH, P. A model of the perceptual and conceptual processes in graph comprehension. *J. of Experimental Psychology: Applied 4*, 2 (1998), 75–100.

[32] CARTER, G. Coherence and time delay estimation. *Proceedings of the IEEE 75*, 2 (Feb. 1987), 236–255.

[33] CHANDLER, P., AND SWELLER, J. Cognitive load theory and the format of instruction. *Cognition and Instruction 8* (1991), 293–332.

[34] CHU, S., KEOGH, E., HARD, D., AND PAZZANI, M. Iterative deepending dynamic time warping for time series. *Second SIAM International Conference on Data Mining* (2002).

[35] CLARK, C., VELTMEYER, M., HAMILTON, R., SIMMS, E., PAUL, R., HERMENS, D., AND GORDON, E. Spontaneous alpha peak frequency predicts working memory performance across the age span. *International Journal of Psychophysiology 43*, 1 (2004), 1–9.

[36] CLEVELAND, W. S., AND MCGILL, R. Graphical perception: Theory, experimentation and the application to the development of graphical methods. *J. of American Statistical Association 79* (1984), 531–554.

[37] COHEN, J. D., PERLSTEIN, W. M., BRAVER, T. S., NYSTROM, L. E., NOLL, D. C., JONIDES, J., AND SMITH, E. E. Temporal dynamics of brain activation during a working memory task. *Nature 386* (1997), 604–608.

[38] CONSTANTINIDIS, C., AND WANG, X.-J. A neural circuit basis for spatial working memory. *The Neuroscientist 10*, 6 (2004), 553–565.

[39] COWAN, N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioural Brain Science 24* (2001), 87–114.

[40] DAVIDSON, S. B., AND FREIRE, J. Provenance and scientific workflows: Challenges and opportunities. In *In Proceedings of ACM SIGMOD* (2008), pp. 1345–1350.

[41] DELORME, A., AND MAKEIG, S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods 134*, 1 (2004), 9–21.

[42] DESMOND, J. E., CHEN, S. H., AND SHIEH, P. B. Cerebellar transcranial magnetic stimulation impairs verbal working memory. *Annals of Neurology 58* (2005), 553–560.

[43] D'ESPOSITO, M., POSTLE, B. R., AND RYPMA, B. Prefrontal cortical contributions to working memory: Evidence from event-related fMRI studies. *Experimental Brain Research 133* (2000), 3–11.

[44] DUNCAN, J., AND OWEN, A. M. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences 23* (2000), 475–483.

[45] ECKHORN, R., BAUER, R., JORDAN, W., BROSCH, M., KRUSE, W., MUNK, M., AND REITBOECK, H. Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological Cybernetics 60*, 2 (1988), 121–130.

[46] ELLKVIST, T., KOOP, D., ANDERSON, E., FREIRE, J., AND SILVA, C. Using provenance to support real-time collaborative design of workflows. In *Provenance and Annotation of Data and Processes*, J. Freire, D. Koop, and L. Moreau, Eds., vol. 5272 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2008, pp. 266–279.

[47] ENGLE, R. W. Working memory capacity as executive attention. *Current Directions in Psychological Science 11* (2002), 19–23.

[48] ESENTHER, A. Instant co-browsing: Lightweight real-time collaborative web browsing, 2002.

[49] FISCHL, B., SALAT, D. H., BUSA, E., ALBERT, M., DIETERICH, M., HASELGROVE, C., VAN DER KOUWE, A., KILLIANY, R., KENNEDY, D., KLAVENESS, S., MONTILLO, A., MAKRIS, N., ROSEN, B., AND DALE, A. M. Whole brain segmentation. *Neuron 33*, 3 (2002), 341–355.

[50] FORD, J., MATHALON, D., WHITFIELD, S., FAUSTMAN, W., AND ROTH, W. Reduced communication between frontal and temporal lobes during talking in schizophrenia. *Biological Psychiatry 51*, 6 (2002), 485–492.

[51] FORD, J. M., KRYSTAL, J. H., AND MATHALON, D. H. Neural synchrony in schizophrenia: From networks to new treatments. *Schizophrenia Bulletin 33* (2007), 848–852.

[52] FREIRE, J., KOOP, D., SANTOS, E., AND SILVA, C. T. Provenance for computational tasks: A survey. *IEEE Computing in Science and Engineering 10*, 3 (Mar 1, 2008), 11–21.

[53] FREIRE, J., SILVA, C. T., CALLAHAN, S. P., SANTOS, E., SCHEIDEGGER, C. E., AND VO, H. T. Managing rapidly-evolving scientific workflows. In *Provenance and Annotation of Data*, L. Moreau and I. Foster, Eds., vol. 4145. Springer Berlin / Heidelberg, 2006, pp. 10–18.

[54] FREIRE, J., SILVA, C. T., CALLAHAN, S. P., SANTOS, E., SCHEIDEGGER, C. E., AND VO, H. T. Managing rapidly-evolving scientific workflows. In *International Provenance and Annotation Workshop (IPAW)* (2006), LNCS 4145, pp. 10–18.

[55] FRISTON, K. J. Schizophrenia and the disconnection hypothesis. *Acta Psychiatrica Scandinavica Supplementum 395* (1999), 68–79.

[56] GABGLIO, A. *Theoria generale della statistica, 2nd Edition*. Milan, 1888.

[57] GEVINS, A., AND SMITH, M. E. Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex 10*, 9 (2000), 829–839.

[58] GEVINS, A., SMITH, M. E., MCEVOY, L., AND YU, D. High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex 7* (1997), 374–385.

[59] G.FUGGETTA, PAVONE, E. F., FIASCHI, A., AND MANGANOTTI, P. Acute modulation of cortical oscillatory activities during short trains of high-frequency repetitive transcranial magnetic stimulation of the human motor cortex: A combined EEG and TMS study. *Human Brain Mapping 29* (2008), 1–13.

[60] GIBSON, A., GAMBLE, M., WOLSTENCROFT, K., OINN, T., AND GOBLE, C. The data playground: An intuitive workflow specification environment. In *E-SCIENCE '07: Proceedings of the Third IEEE International Conference on e-Science and Grid Computing* (Washington, DC, USA, 2007), IEEE Computer Society, pp. 59–68.

[61] GOBLE, C. A., AND ROURE, D. C. D. myExperiment: Social networking for workflow-using e-scientists. In *WORKS '07: Proceedings of the 2nd workshop on Workflows in Support of Large-Scale Science* (New York, NY, USA, 2007), ACM, pp. 1–2.

[62] GOEL, V., AND GRAFMAN, J. Are the frontal lobes implicated in planning functions? Interpreting data from The Tower of Hanoi. *Neuropsychologia 33* (1995), 623–642.

[63] GONZALEZ, R. C., AND WOODS, R. E. *Digital Image Processing, 3rd Edition*. Prentice Hall, 2007.

[64] GREEN, M., NUECHTERLEIN, K., GOLD, J., BARCH, D., COHEN, J., ESSOCK, S., FENTON, W., FRESE, F., GOLDBERG, T., HEATON, R., KEEFE, R., KEM, R., KRAEMER, H., STOVER, E., WEINBERGER, D., ZALCMAN, S., AND MARDER, S. Approaching a consensus cognitive battery for clinical trials in schizophrenia. *NIMH-MATRIC Conference to Select Cognitive Domains and Test Criteria 56* (2004), 301–330.

[65] GRISARU, N., AMIR, M., COHEN, H., AND KAPLAN, Z. Effect of transcranial magnetic stimulation in posttraumatic stress disorder: A preliminary study. *Biological Psychiatry 44* (1998), 52–55.

[66] HAMIDI, M., TONONI, G., AND POSTLE, B. R. Evaluating the role of prefrontal and parietal cortices in memory-guided response with repetitive transcranial magnetic stimulation. *Neuropsychologia 47* (2009), 295–302.

[67] HARVEY, P., GREEN, M., MCGURK, S., AND MELTZER, H. Changes in cognitive functioning with risperidone and olanzapine treatment: A large-scale, double-blind, randomized study. *Psychopharmacology 169*, 3–4 (2003), 404–411.

[68] HERWIG, U., SATRAPI, P., AND SCHNFELDT-LECUONA, C. Using the international 10-20 eeg system for positioning of transcranial magnetic stimulation. *Brain Topography 16* (2003), 95–99. 10.1023/B:BRAT.0000006333.93597.9d.

[69] HINTZE, J. L., AND NELSON, R. D. Violin plots: A box plot-density trace synergism. *The American Statistician 52*, 2 (1998), 181–184.

[70] HITT, J. M., KRING, J. P., DASKAROLIS, E., MORRIS, C., AND MOULOUA, M. Assessing mental workload with subjective measures: An analytical review of the nasa-tlx index since its inception. *Human Factors and Ergonimics Society Annual Meeting 43* (1999), 1404–1404.

[71] HOFFMAN, R. E., GUEORGUIEVA, R., HAWKINS, K. A., VARANKO, M., BOUTROS, N. N., AND WU, Y. T. Temporoparietal transcranial magnetic stimulation of auditory hallucinations: Safety, efficacy and moderators in a fifty patient sample. *Biological Psychiatry 58* (2005), 97–104.

[72] HOWE, B., LAWSON, P., BELLINGER, R., ANDERSON, E. W., SANTOS, E., FREIRE, J., SCHEIDEGGER, C., BAPTISTA, A., AND SILVA, C. End-to-end eScience: Integrating workflow, query, visualization, and provenance at an ocean observatory. In *In Proceedings of Fourth IEEE International Conference on eScience* (2008), pp. 127–134.

[73] HULL, D., WOLSTENROFT, K., STEVENS, R., GOBLE, C., POCOCK, M., LI, P., AND OINN, T. Taverna: A tool for building and running workflows of services. *Nucleic Acids Research 34* (2006), 729–732.

[74] HUNTER, J. D. Matplotlib: A 2d graphics environment. *Computing In Science & Engineering 9*, 3 (May-Jun 2007), 90–95.

[75] HYVÄRINEN, A., HOYER, P. O., AND INKI, M. Independent component analysis. *Neural Computing Surveys 2* (2001).

[76] JUNG, T.-P., MAKEIG, S., WESTERFIELD, M., TOWNSEND, J., COURCHESNE, E., AND SEJNOWSKI, T. Analysis and visualization of single-trial event-related potentials. *Human Brain Mapping 14* (2001), 166–185.

[77] KLIMESCH, W. EEG-alpha rhythms and memory processes. *International Journal of Psychophysiology 26* (1997), 319–340.

[78] KLIMESCH, W. EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews 29* (1999), 169–195.

[79] KLIMESCH, W., SAUSENG, P., AND GERLOFF, C. Enhancing cognitive performance with repetitive transcranial magnetic stimulation at human individual alpha frequency. *European Journal of Neuroscience 17* (2003), 1129–1133.

[80] KLINGNER, J., TVERSKY, B., AND HANRAHAN, P. Effects of visual and verbal presentation on cognitive load in vigilance, memory and arithmetic tasks. *J. of Psyhophysiology 48*, 3 (2011), 323–332.

[81] KOCH, G., OLIVERI, M., TORRIERO, S., CARLESIMO, G. A., TURRIZIANI, P., AND CALTAGIRONE, C. rTMS evidence of different delay and decision processes in a fronto-parietal neuronal network activated during spatial working memory. *Neuroimage 24* (2005), 34–39.

[82] KOENIG, T., MARTI-LOPEZ, F., AND VALDES-SOSA, P. Topographic time-frequency decomposition of the EEG. *NeuroImage 14* (August 2001), 383–390.

[83] KOHLER, S., PAUS, T., BUCKNER, R. L., AND MILNER, B. Effects of left inferior prefrontal stimulation on episodic memory formation: A two-stage fMRI–rTMS study. *Journal of Cognitive Neuroscience 16* (2004), 178–188.

[84] KOOP, D., SANTOS, E., SCHEIDEGGER, C. E., VO, H. T., SILVA, C. T., AND FREIRE, J. VisTrails. In *Architecture of Open-Source Applications*, G. Wilson, Ed. Jun 1 2011. To Appear.

[85] KOSARA, R., HEALEY, C. G., INTERRANTE, V., LAIDLAW, D. H., AND WARE, C. Thoughts on user studies: Why, how and when. *IEEE Computer Graphics and Applications 23*, 4 (2003), 20–25.

[86] KOSARA, R., MIKSCH, S., AND HAUSER, H. Semantic depth of field. *Proceedings of IEEE INFOVIS* (2001), 97–104.

[87] KRAEPELIN, E. Compendium der psychiatrie. *A. Abel, 1883. Psychiatrie: Ein Lehrbuch fur Studierende und rtze.4th edition,. 1883.*.

[88] LAIDLAW, D. H., KIRBY, R. M., JACKSON, C. D., DAVIDSON, J. S., MILLER, T. S., DA SILVA, M., WARREN, W. H., AND TARR, M. J. Comparing 2d vector field visualization methods: A user study. *IEEE Transactions on Visualization and Computer Graphics 11*, 2 (2005), 59–70.

[89] LEIBERG, S., LUTZENBERGER, W., AND KAISER, J. Effects of memory load on cortical oscillatory activity during auditory pattern working memory. *Brain Research 1120* (2006), 131–140.

[90] LIKERT, R. A technique for the measurement of attitudes. *Archives of Psychology 140* (1932), 1–55.

[91] LIN, J., KEOGH, E., AND LONARDI, S. Visualizing and discovering non-trivial patterns in large time series databases. *Information Visualization 4*, 2 (2005), 61–82.

[92] LIN, J., KEOGH, E., LONARDI, S., LANKFORD, J. P., AND NYSTROM, D. M. Viztree: A tool for visually mining and monitoring massive time series databases. In *VLDB 2004: Proceedings of the Thirtieth International Conference on Very Large Data Bases* (2004), VLDB Endowment, pp. 1269–1272.

[93] LINKENKAER-HANSEN, K., MONTO, S., RYTSÄLÄ, H., SUOMINEN, K., ISOMETSÄ, E., AND KÄHKÖNEN, S. Breakdown of long-range temporal correlations in theta oscillations in patients with major depressive disorder. *Journal of Neuroscience 25* (2005), 10131–10137.

[94] LUBER, B., KINNUNEN, L. H., RAKITIN, B. C., ELLSASSER, R., STERN, Y., AND LISANBY, S. H. Facilitation of performance in a working memory task with rTMS stimulation of the precuneus: Frequency- and time-dependent effects. *Brain Research 1128* (2007), 120–129.

[95] LUBER, B., STANFORD, A. D., BULOW, P., NHUYEN, T., RAKITIN, B. C., AND HABECK, C. Remediation of sleep-deprivation induced working memory impairment with fMRI-guided transcranial magnetic stimulation. *Cerebral Cortex 18* (2008), 2077–2085.

[96] MALLAT, S. A theory for multiresolution signal decomposition: The Wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence 11*, 7 (Jul 1989), 674–693.

[97] MARKS, J., ANDALMAN, B., BEARDSLEY, P. A., FREEMAN, W., GIBSON, S., HODGINS, J., KANG, T., MIRTICH, B., PFISTER, H., RUML, W., RYALL, K., SEIMS, J., AND SHIEBER, S. Design galleries: A general approach to setting parameters for computer graphics and animation. In *SIGGRAPH '97: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1997), ACM Press/Addison-Wesley Publishing Co., pp. 389–400.

[98] MAROIS, R., AND IVANOFF, J. Capacity limits of information processing in the brain. *Trends in Cognitive Sciences 9* (2005), 269–305.

[99] MCGURK, S., HARVEY, P., LAPUGLIA, R., AND MARDER, J. Cognitive and symptom predictors of work outcomes for clients with schizophrenia in supported employment. *Psychiatric Services 54*, 8 (2003), 1129–1135.

[100] MCNAMES, J., BASSALE, J., ABOY, M., CRESPO, C., AND GOLDSTEIN, B. Techniques for the visualization of nonstationary biomedical signals. In *Proceedings of the 16th International EURASIP Conference BIOSIGNAL* (2002), vol. 16, pp. 42–45.

[101] MEYER, Y. *Wavelets - Algorithms and applications.* Society for Industrial and Applied Mathematics, 1993.

[102] MISHARA, A., AND GOLDBERG, T. A meta-analysis and critical review of the effects of conventional neuroleptic treatment on cognition in schizophrenia: Opening a closed book. *Biological Psychiatry 55*, 10 (2004), 1013–1022.

[103] MOTTAGHY, F. M., GANGITANO, M., KRAUSE, B. J., AND PASCUAL-LEONE, A. Chronometry of parietal and prefrontal activations in verbal working memory revealed by transcranial magnetic stimulation. *Neuroimage 18* (2003), 167–170.

[104] MOTTAGHY, F. M., KRAUSE, B. J., KEMNA, L. J., TÖPPER, R., TELLMANN, L., AND BEU, M. Modulation of the neuronal circuitry subserving working memory in healthy human subjects by repetitive transcranial magnetic stimulation. *Neuroscience Letters 280* (2000), 167–170.

[105] MOTTAGHY, F. M., SPARING, R., AND TOPPER, R. Enhancing picture naming with transcranial magnetic stimulation. *Behavioral Neurology 17* (2006), 177–186.

[106] MULL, B. R., AND SEYAL, M. Transcranial magnetic stimulation of left prefrontal cortex impairs working memory. *Clinical Neurophysiology 112* (2001), 1672–1675.

[107] NIEDERMEYER, E., AND DA SILVA, F. L. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.* Lippincott Williams and Wilkins, 2005.

[108] NUWER, M. R. Quantitative EEG: Techniques and problems of frequency analysis and topographic mapping. *Journal of Clinical Neurophysiology 5*, 1 (1988).

[109] OOSTENVELD, R., FRIES, P., MARIS, E., AND SCHOFFELEN, J.-M. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience 2011*, 156869 (2011).

[110] O'REARDON, J. P., SOLVASON, H. B., JANICAK, P. G., SAMPSON, S., ISENBERG, K. E., AND NAHAS, Z. Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: A multisite randomized controlled trial. *Biological Psychiatry 62* (2007), 1208–1216.

[111] OSAKA, N., OTSUKA, Y., HIROSE, N., IKEDA, T., MIMA, T., AND FUKUYAMA, H. Transcranial magnetic stimulation (tms) applied to left dorsolateral prefrontal cortex disrupts verbal working memory performance in humans. *Neuroscience Letters 418* (2007), 232–235.

[112] OSIPOVA, D., AHVENINEN, J., JENSEN, O., YLIKOSKI, A., AND PEKKONEN, E. Altered generation of spontaneous oscillations in Alzheimer's disease. *Neuroimage 27* (2005), 835–841.

[113] OSUCH, E. A., BENSON, B. E., LUCKENBAUGH, D. A., GERACI, M., POST, R. M., AND MCCANN, U. Repetitive TMS combined with exposure therapy for PTSD: A preliminary study. *Journal of Anxiety Disorder* (2009), 54–59.

[114] PAAS, F., RENKL, A., AND SWELLER, J. Cognitive load theory and instructional design: Recent developments. *Educational Psychologist 38*, 1 (2003), 1–4.

[115] PIROLLI, P., CARD, S. K., AND VAN DER WEGE, M. M. Visual information foraging in a focus + context visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2001), CHI '01, ACM, pp. 506–513.

[116] PLEWNIA, C., FILK, A. J., SOEKADAR, S. R., ARFELLAR, C., HUBER, H. S., AND SAUSENG, P. Enhancement of long-range EEG coherence by synchronous bifocal transcranial magnetic stimulation. *European Journal of Neuroscience 27* (2008), 1577–1583.

[117] POSTLE, B. R., FERRARELLI, F., HAMIDI, M., FEREDOES, E., MASSIMINI, M., AND PETERSON, M. Repetitive transcranial magnetic stimulation dissociates working memory manipulation from retention functions in the prefrontal but not posterior parietal cortex. *Journal of Cognitive Neuroscience 18* (2006), 1712–1722.

[118] POTTER, K., KNISS, J., RIESENFELD, R., AND JOHNSON, C. R. Visualizing summary statistics and uncertainty. *Computer Graphics Forum 29*, 3 (2010), 823–831.

[119] PRESTON, G., AND WEINBERGER, D. Intermediate phenotypes in schizophrenia: A selective review. *Dialogues in Clinical Neuroscience 7*, 2 (2005), 165–179.

[120] PRESTON, G. A., ANDERSON, E. W., GOLDBERG, T., WASSERMANN, E., WEINBERGER, D., AND SILVA, C. T. Effects of 10Hz rTMS on the neural efficiency of working memory. *Journal of Cognitive Neuroscience* (2009).

[121] REITINGER, B., BORNIK, A., BEICHEL, R., AND SCHMALSTIEG, D. Liver surgery planning using virtual reality. *IEEE Computer Graphics and Applications 26* (2006), 36–47.

[122] RICHARDSON, T., STAFFORD-FRASER, Q., WOOD, K. R., AND HOPPER, A. Virtual network computing. *IEEE Internet Computing 2*, 1 (1998), 33–38.

[123] ROSENHECK, R., BINGHAM, S., LIU-MARES, W., COLLINS, J., WARREN, S., AND LESLIE, D. Effectiveness and cost of olanzapine and haloperidol in the treatment of schizophrenia: A randomized controlled trial. *JAMA 290*, 20 (2003), 2693–2702.

[124] ROUNIS, E., YARROW, K., AND ROTHWELL, J. C. Effects of rTMS conditioning over the fronto-parietal network on motor versus visual attention. *Journal of Cognitive Neuroscience 19* (2007), 513–524.

[125] ROWE, J. B., TONI, I., JOSEPHS, O., FRACKOWIAK, R. S. J., AND PASSINGHAM, R. E. The prefrontal cortex: Response selection or maintenance within working memory. *Science 288*, 5471 (2000), 1656–1660.

[126] ROWEIS, S. T., AND SAUL, L. K. Nonlinear dimensionality reduction by locally linear embedding. *Science 290*, 22 (200), 2323–2326.

[127] RYPMA, B., BERGER, J. S., AND D'ESPOSITO, M. The influence of working memory demand on subject performance on prefrontal cortical activity. *Journal of Cognitive Neuroscience 14* (2002), 721–731.

[128] RYPMA, B., BERGER, J. S., PRABHAKARAN, V., BLY, B. M., KIMBERG, D. Y., AND BISWAL, B. B. Neural correlates of cognitive efficiency. *Neuroimage 33* (2006), 969–979.

[129] RYPMA, B., PRABHAKARAN, V., DESMOND, J. E., AND GABRIELI, J. D. Age differences in prefrontal cortical activity in working memory. *Psychology and Aging 16* (2001), 371–384.

[130] SALENIUS, S., PORTIN, K., KAJOLA, M., SALMELIN, R., AND HARI, R. Cortical control of human mononeuron firing during isometric contraction. *J. of Neurophysiology 77*, 6 (1997), 3401–3405.

[131] SANTOS, E., KOOP, D., VO, H. T., ANDERSON, E. W., FREIRE, J., AND SILVA, C. Using workflow medleys to streamline exploratory tasks. In *Proceedings of the 21st International Conference on Scientific and Statistical Database Management* (Berlin, Heidelberg, 2009), SSDBM 2009, Springer-Verlag, pp. 292–301.

[132] SANTOS, E., LINS, L., AHRENS, J., FREIRE, J., AND SILVA, C. Vismashup: Streamlining the creation of custom visualization applications. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 1539–1546.

[133] SANYAL, J., ZHANG, S., BHATTACHARYA, G., AMBURN, P., AND MOORHEAD, R. J. A user study to compare four uncertainty visualization methods for 1d and 2d datasets. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 1209–1218.

[134] SARTER, N. B., AND WOODS, D. D. Pilot interaction with cockpit automation ii: An experimental study of pilots' model and awareness of the flight management system. *Int'l J. of Aviation Psychology 4*, 1 (1994), 1–28.

[135] SAUSENG, P., AND KLIMESCH, W. EEG alpha synchronization and functional coupling during top-down processing in a working memory task. *Human Brain Mapping 26*, 1 (2005), 148–155.

[136] SCHEIDEGGER, C., VO, H., KOOP, D., FREIRE, J., AND SILVA, C. Querying and creating visualizations by analogy. *IEEE Transactions on Visualization and Computer Graphics 13* (2007), 1560–1567.

[137] SCHIFF, S. J., ALDROUBI, A., UNSER, M., AND SATO, S. Fast wavelet transformation of EEG. *Electroencephalography and Clinical Nuerophysiology 91*, 6 (1994), 442–455.

[138] SCHROEDER, W., AND LORENSON, B. *Visualization Toolkit: An Object-Oriented Approach to 3-D Graphics*, 1st ed. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1996.

[139] SELEMON, L., AND GOLDMAN-RAKIC, P. Elevated neuronal density in prefrontal area 46 in brains from schizophrenic patients: Application of a three-dimensional, stereologic counting method. *Journal of Computational Neurology 392*, 3 (1998), 402–412.

[140] SEUFERT, T., JÄNEN, I., AND BRÜNKEN, R. The impact of intrinsic cognitive load on the effectiveness of graphical help for coherence information. *Computers in Human Behavior 23* (2007), 1055–1071.

[141] SHI, Y., RUIZ, N., TAIB, R., CHOI, E., AND CHEN, F. Galvanic skin response (GSR) as an index of cognitive load. In *CHI '07 extended abstracts* (2007), ACM, pp. 2651–2656.

[142] SILVA, C. T., ANDERSON, E. W., SANTOS, E., AND FREIRE, J. Using vistrails and provenance for teaching scientific visualization. *Computer Graphics Forum 29*, 3 (2010).

[143] SIMMHAN, Y., PLALE, B., AND GANNON, D. A survey of data provenance in e-science. *SIGMOD Record 34*, 3 (2005), 31–36.

[144] SINKKONEN, J., TIITNEN, H., AND NÄÄTÄNEN, R. Gabor filters: An informative way for analyzing event-related brain activity. *Journal of Neuroscience Methods 56*, 1 (1995), 99–104.

[145] STERNBERG, S. High-speed scanning in human memory. *Science 153*, 3736 (1966), 652–654.

[146] STERNBERG, S. Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist 57* (1969), 421–457.

[147] STOCKWELL, R. G. A basis for efficient representation of the s-transform. *Digital Signal Processing 17*, 1 (2007), 371–393.

[148] SWELLER, J. *The Cambridge handbook of multimedia learning.* Cambridge University Press, 2005, ch. Implications of cognitive load theory for multimedia learning, pp. 19–30.

[149] TALAIRACH, J., AND TOURNOUX, P. *Co-Planar Stereotaxic Atlas of the Human Brain - 3-Dimensional Proportional System: An Approach to Cerebral Imaging*, 1st ed. Thieme, 1988.

[150] TEN CAAT, M., MAURITS, N. M., AND ROERDINK, J. Functional unit maps for data-driven visualization of high-density EEG coherence. In *Proc. Eurographics/IEEE VGTC Symposium on Visualization (EuroVis)* (2007), pp. 259–266.

[151] TEN CAAT, M., MAURITS, N. M., AND ROERDINK, J. B. Design and evaluation of tiled parallel coordinate visualization of multichannel EEG data. *IEEE Transactions on Visualization and Computer Graphics 13*, 1 (2007), 70–79.

[152] TENENBAUM, J. B., DE SILVA, V., AND LANGFORD, J. C. A global geometric framework for nonlinear dimensionality reduction. *Science 290*, 22 (200), 2319–2323.

[153] TORY, M., AND MÖLLER, T. Human factors in visualization research. *IEEE Transactions on Visualization and Computer Graphics 10*, 1 (2004), 72–84.

[154] TUFTE, E. R. *The Visual Display of Quantitative Information*. Graphics Press, 1983.

[155] TUKEY, J. W. *Exploratory Data Analysis*. Addison-Wesley, 1977.

[156] UHLHAAS, P. J., AND SINGER, W. What do distrubances in neural synchrony tell us about autism? *Biological Psychiatry 62* (2007), 190–191.

[157] VOGT, S., BUCCINO, G., WOHLSCHLAGER, A. M., CANESSA, N., SHAH, N. J., AND ZILLES, K. Prefrontal involvement in imitation learning of hand actions: Effects of practice and experience. *Neuroimage 37* (2007), 1371–1383.

[158] WANG, H., BRODLIE, K., HANDLEY, J., AND WOOD, J. Service-oriented approach to collaborative visualization. In *Proceedings of UK e-Science All Hands Meeting 2006* (2006), National e-Science Centre, pp. 241–248.

[159] WASSERMANN, E. M. Risk and safety of repetitive transcranial magnetic stimulation: Report and suggested guidelines from the international workshop on the safety of repetitive transcranial magnetic stimulation. *Electroencephalography and Clinical Neurophysiology 108* (1998), 1–16.

[160] WATTENBERG, M. Arc diagrams: Visualizing structure in strings. *In Proceedings of the IEEE Symposium on Information Visualization* (2002).

[161] WEBER, M., ALEXA, M., AND MULLER, W. Visualizing time-series on spirals. *IEEE Symposium on Information Visualization, 2001* (2001), 7–13.

[162] WOOD, J., WRIGHT, H., AND BRODLIE, K. Collaborative visualization. In *VIS '97: Proceedings of the 8th Conference on Visualization '97* (Los Alamitos, CA, USA, 1997), IEEE Computer Society Press, pp. 253–ff.

[163] Yahoo! Pipes. `http://pipes.yahoo.com` [10 March 2008].

[164] YANKOV, D., KEOGH, E., LONARDI, S., AND FU, A. W. Dot plots for time series analysis. In *Proceedings of IEEE International Conference on Tools with Artificial Intelligence (ICTAI'05)* (2005), pp. 159–168.

[165] YENER, G. G., GÜNTEKIN, B., ONIZ, A., AND BARAR, E. Increased frontal phase-locking of event-related theta oscillations in alzheimer patients treated with cholinesterase inhibitors. *International Journal of Psychophysiology 64* (2007), 46–52.