Developing and Evaluating a Geographic Information Dashboard to Improve Spatial Task Performance

by

Amariah Joanne Fischer

B.S., University of Nebraska-Lincoln, 2016

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF ARTS

Department of Geography
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Dr. J.M. Shawn Hutchinson

Copyright

© Amariah Fischer 2018.

Abstract

Information dashboards are decision-support tools that pull data from multiple sources and display those data on a single screen. Information dashboards are becoming common in fields such as medicine, computer science, and business, given their perceived ability to facilitate faster and more accurate assessments by users. However, there is very little peer-reviewed research on information dashboards that support this assumption. This research focuses on applying the concept of an information dashboard visualization within the spatial sciences and evaluating the effectiveness of a geographic information dashboard, or GID, on improving user performance related to spatial thinking tasks. A review of literature from multiple disciplines highlights what is, and what is not, understood about dashboard visualizations. Borrowing from ideas such as Cognitive Fit Theory and past work in evaluating the effectiveness of map animations, an appropriate method for evaluating the GID is proposed.

A Web-based GID and an alternative "tabbed" visualization were developed using the R Shiny package to support an analysis of grassland vegetation development for a site located in northeastern Kansas. A controlled experiment was conducted using a survey completed by volunteer student participants who responded to a series of benchmark tasks related to the interpretation of 6 related maps and graphs. Data for three dependent variables (task completion time, task response accuracy, and an integrative measure of performance accounting for both time and accuracy) were collected directly from the survey or post-survey grading of responses. Three independent variables and their impact on spatial task performance were analyzed, including the type of visualization, assessed spatial thinking ability, and cognitive task type.

Results showed that participants using the GID completed the benchmark tasks faster and more accurately, but that a users' spatial thinking ability had the most significant influence on performance regardless of visualization. Evidence was found to support the idea that the GID improved spatial thinking performance, especially for users with more experience in spatial reasoning, and that the GID format may improve user performance beyond what is expected based on an independent assessment of spatial thinking ability.

Table of Contents

List of Figures

List of Tables

Acknowledgments

Firstly, I would like to acknowledge the significant contributions of my graduate committee members, Dr. Shawn Hutchinson, Dr. Lisa Harrington, and Dr. Brent Chamberlain. Without their advice, insights, and guidance I would never have been able to complete my master's thesis. Secondly, I would like to thank Dr. Rebecca Buller, Dr. David Wishart, and Dr. Clark Archer for sparking my interest in geography as an undergraduate at the University of Nebraska-Lincoln. Lastly, I would like to thank the forty-one K-State students who participated in my experiment, for without them, none of this was possible.

Dedication

I would like to dedicate this thesis to my very large and rowdy family. While they are not as enthusiastic about my research as I am, they are always supportive and encouraging. I would also like to dedicate this thesis to my friends at K-State, especially Abbey, Brenna, and Matt Brooks, for spending excessive amounts of time in my office, forcing me to take a break from school work and very likely preserving my sanity.

Chapter 1 - Introduction

Visualization has been described as a "method of computing" and "seeing the unseen" as it can lead to "profound and unexpected insights" (McCormick, DeFanti, and Brown 1987). However, as visualization has progressed and expanded outside of the field of computer science, other perspectives on visualization have developed. Visualization today is defined as the representation of data in a graphical form for the purpose of gaining a new understanding of or revealing new information about a phenomenon that would otherwise be hidden (Tobón, 2005). For this reason, visualization research can be found in many disciplines such as computer science, mathematics, science education, and geography. The research area that comprises the intersection of visualization and geography is known as geographic visualization or geovisualization. This thesis is a geographic visualization study focused specifically on the information dashboard visualization.

The "information dashboard" has been gaining momentum in medicine, computer science, and business intelligence over the last three decades due to its usefulness as a decision-support tool. Information dashboards are designed to connect data from various sources needed to address a specific problem by displaying those data comprehensively and on a single computer screen. By designing the visualization in such a way, a user is believed to achieve a better understanding of the data as a whole and in a shorter amount of time, leading to more accurate assessments. However, as Yigitbasioglu and Velcu (2012) point out, there is very little peer-reviewed literature on the effectiveness of information dashboards to support the claim that they do, indeed, lead to faster, more accurate decisions. While Few (2006) makes logical arguments for specific design techniques to promote effective dashboard visualizations, his design suggestions are not supported by research or experiments that robustly test his theories.

As geography is a discipline that addresses a multitude of complex, real-world problems that require the gathering of spatial, spatiotemporal, and aspatial data from a variety of sources, it is particularly well-suited as an application domain within which information dashboard visualizations may be studied. Therefore, the goals of this research were to 1) apply the dashboard visualization to geography by creating a geographic information dashboard (GID) and to 2) evaluate the GID for effectiveness to contribute to the small amount of scholarly research on information dashboards.

To achieve these goals, Chapter 2 of this thesis consists of an in-depth literature review explaining how the development and testing of a GID would fill multiple gaps in the literature, making the argument for why it is worth investing time and resources in this visualization. Following the argument for why the GID is a necessary development, research from multiple disciplines is used to develop the framework of a rigorous method for evaluating the GID.

Chapter 3 provides a detailed account of the methods used to create and evaluate the GID, while Chapter 4 covers the results of the evaluation. Finally, in Chapter 5 the results, limitations, and broader impacts of this study are discussed

Chapter 2 - Literature Review

2.1 Introduction

Maps have long been the medium for visualizing spatial data, from navigational purposes in the 15th and 16th centuries to visualizing socio-economic inequalities today. Advances in geographic visualization research have allowed researchers to expand beyond traditional cartographic approaches and explore other mediums for visualizing spatial data and phenomena. There is great interest in generating new geographic visualizations because they have the ability to uncover patterns in data that may have been otherwise hidden. The purpose of this chapter is to propose a new geographic visualization, or geoviz, in the form of a geographic information dashboard (GID). While technological advances have made it much easier to make new and interesting visualizations, understanding why a visualization is needed and how it can be tested are vital steps in the research process. This chapter will establish why a GID is needed by exploring important gaps in current visualization research, then will propose an approach by which a GID may be tested using scholarly methods derived from a variety of disciplines.

2.2 Why the scientific community needs a GID

The first section of this paper focuses on why a GID is a geographic visualization in which it would be valuable to invest time and resources. To accomplish this, three areas of study that will benefit from GID research are identified: information dashboards, geographic visualization development, and geographic visualization testing. The conceptual process for establishing specific research gaps the GID will fill is presented below (Figure 1). The top row of the figure identifies what is known from prior research and peer-reviewed literature, while the bottom row addresses gaps. The following section will individually address each of the specified research areas to not only show that these gaps exist but to also show how the development and evaluation of the GID will begin to close gaps in the literature.

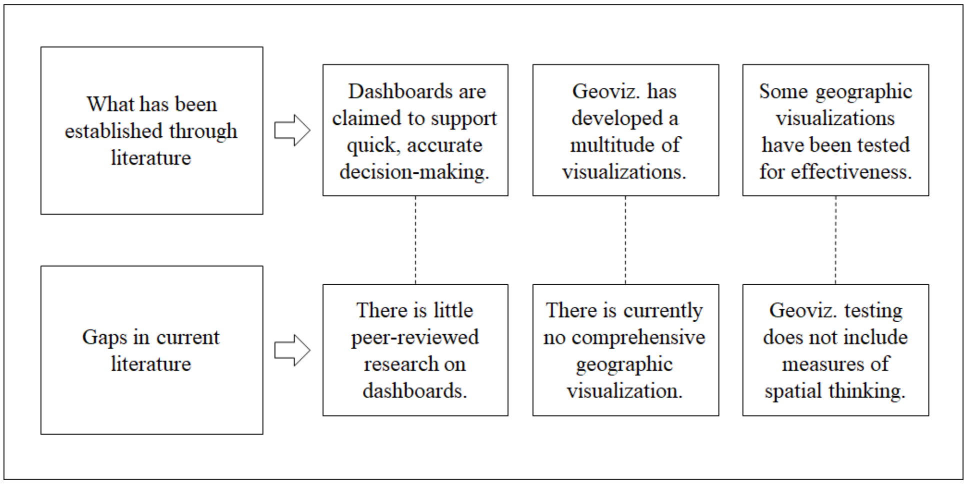


Figure 1 Conceptual process identifying what is known in the literature and the gaps that can be addressed through GID research.

2.2.1 Information Dashboards

Information dashboards first became popular in the 1990s as technology and better methods for data warehousing were developed. Information dashboards are commonly developed within computer science or engineering but are applied in the medical field to help healthcare professionals easily monitor patients, and in the business sector to help owners and upper management quickly evaluate the financial status of a company or corporation (Figure 2). More specifically, though, an information dashboard can be described as:

> A visual and interactive performance management tool that displays on a single screen the most important information to achieve one or several individuals and/or organizational objectives, allowing the user to identify, explore, and communicate problems areas that need corrective action. (Yigitbasioglu and Velcu 2011)

This definition addresses key advantages of the information dashboard visualization, while also exposing a few concerns. Information dashboards are believed to allow users to quickly and easily identify problems revealed in the data by presenting key performance indicators in a consciously chosen, understandable format. Based on this definition, it can be seen how information dashboards can be beneficially applied to disciplines that address multi-faceted problems that require the collection of data from various sources, such as geography. However, this definition also requires the dashboard to allow for user exploration. Interactive dashboards are common in business applications, but in medicine dashboards are typically non-interactive, causing dashboard literature stemming from medicine to be less transferable to other disciplines. The final concern stemming from this definition is that, regardless of application, dashboards are claimed to lead to quick, accurate conclusions, but little evidence exists to support this claim.

Figure 2 Example of a business-oriented dashboard design from Klipfolio (www.klipfolio.com).

To better illustrate this point, a simple non-statistical meta-analysis was executed using peer-reviewed articles from 1985 to 2018 containing the term “dashboard” in the article title. Articles were collected from three databases: SCOPUS, Web of Science, and Geobase. SCOPUS and Web of Science were selected for their diverse range of scientific subject areas where dashboard applications might be found. Geobase was selected for its emphasis on Earth sciences, particularly human and physical geography.

Results of this analysis show that there is little peer-reviewed research on dashboards in general, whether the focus is the development or evaluation of a dashboard (Figure 3). The maximum number of publications can be found in SCOPUS in 2017, at 61, and while there is an increasing trend in the number of publications in the last three decades, overall, there are still very few publications on the subject. Secondly, it should be noted that there are only 13 publications found in Geobase for the entire time frame, emphasizing the lack of dashboard literature in geography and allied Earth science disciplines. Furthermore, the subject areas where the majority of dashboard literature can be found are medicine, computer science, and engineering (Figure 3). This is significant because, as previously noted, the dashboards within the medical professional serve different purposes than dashboards in other disciplines, essentially eliminating the usefulness of a large portion of existing dashboard literature for this specific application of the information dashboard visualization.

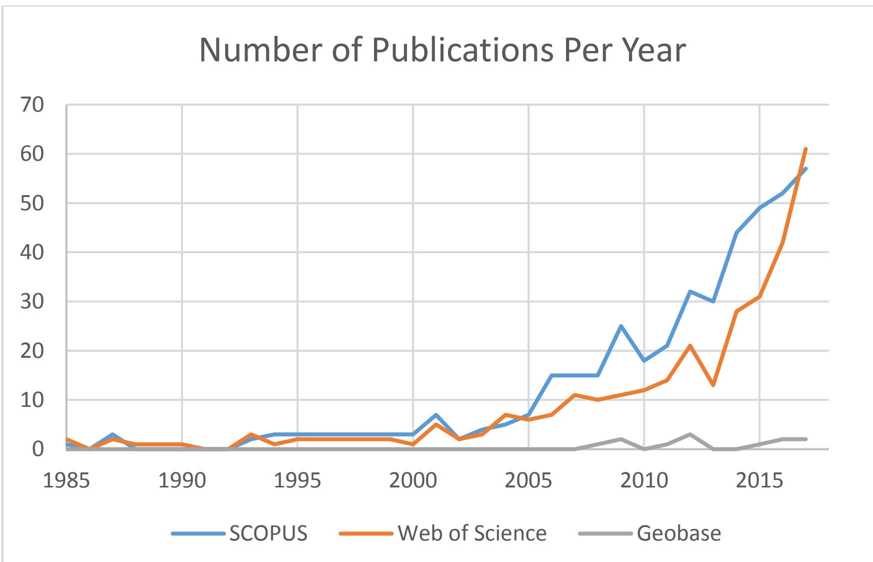


Figure 3 Publication trends in dashboard research from 1985 to 2018; total number of publications per database are as follows: SCOPUS, 430; Web of Science, 311; Geobase, 9.

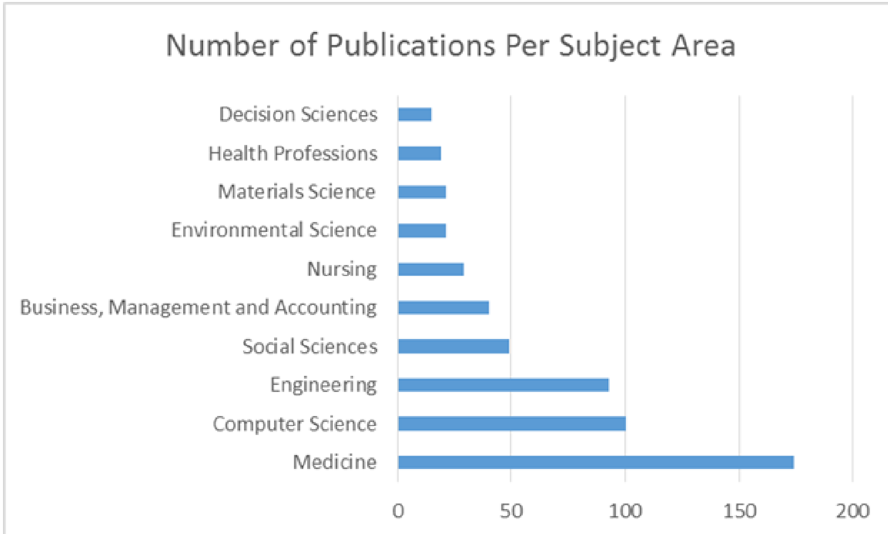


Figure 4 The top ten subject areas where dashboard publications are found using the 430 publications recorded in the SCOPUS database between 1985 and 2018.

An additional meta-analysis was conducted to identify how many of the articles found containing the term "dashboard" in their titles were focused on evaluating the dashboard visualization, as opposed to being centered on the development of a dashboard. This was completed by searching for articles using the term "dashboard" in conjunction with the term "evaluation" or "validation" in the article title. Using the same databases, the results showed that only 15 articles in SCOPUS, 9 articles in Web of Science, and 0 articles in Geobase focus on the evaluation of the dashboard visualization. Additionally, within the 15 articles found in the SCOPUS database, nearly half were from the medical field (Figure 5).

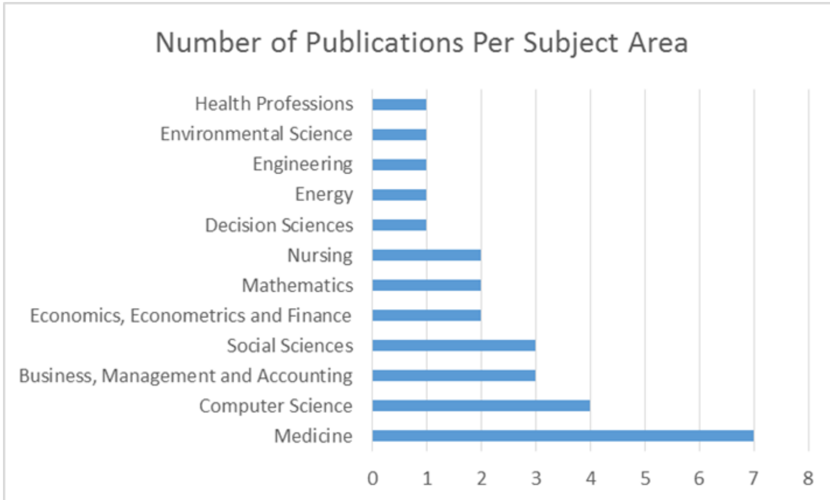


Figure 5 Number of publications per subject area for the fifteen articles on dashboard evaluation found in the SCOPUS database.

This meta-analysis indicates that the dashboard presentation is limited to a small number of disciplines and echoes what has already been noted by Yigitbasioglu and Velcu (2012): while information dashboards are a visualization with great potential to be extremely useful decision-support tools, there is not enough scholarly literature on the subject to inform the design or evaluation of effective dashboards. The development and evaluation of a GID will not only apply the dashboard format to a different discipline, geography, but will also add to our knowledge of the effectiveness of dashboards as a visualization.

2.2.2 Development of New Geographic Visualizations

In addition to contributing to information on dashboard research, GIDs will also address a gap in research by focusing on the development of a new geographic visualization tool. This section illustrates how the GID fills a gap in geovisualization research by being the first comprehensive visualization specifically engineered for effective decision-making. This is accomplished, first, by clarifying the purpose and definition of geovisualization as an extension of visualization. Then, by reviewing the progress in geographic visualization development, it can be seen why the GID is the next logical geographic visualization to be developed.

Geographic visualization has many different definitions provided by multiple authors. One of the most commonly cited definitions is provided by MacEachren and Kraak (2001) who refer to geovisualization as the visual representation of spatial data for the purpose of interactive exploration, hypothesis formation, and analysis. Building from this definition, Dykes, MacEachren, and Kraak (2005) note geographic visualization can lead to new insights and aid in decision making. Lastly, Dodge, McDerby, and Turner (2008) require geographic visualization facilitate spatial understanding. Surveying definitions of geographic visualization reveal a variety of functions, but all should lead to the same result: new methods for representing spatial data so the viewer can best interpret spatial phenomena and processes.

Geographic visualization research first began to flourish in the 1970s and 1980s with the development of Geographic Information Systems (GIS) and the continued dominance of quantitative methods (Hermansen 2010). During this time, geographers became more concerned with user needs throughout the mapping process which lead researchers to focus on how map users interpret different cartographic designs. In the early 1970s, Flannery (1971) discovered that people can correctly estimate length, but tend to underestimate area and volume. This research had implications for the absolute scaling of proportional symbols on a map. While absolute scaling represents data accurately, the map user's interpretation of the data is incorrect because of perceptual scaling. Appearance compensation is used to counteract perceptual scaling and represents one of the ways creating new presentations of spatial data can enhance a user's understanding of spatial phenomena.

Other geovisualization resources aided in choosing appropriate map projections and providing appropriate contrast among map elements (Dent, Torguson, and Hodler 2009; Brewer 2005). Much of this progress in geographic visualization stemmed not from a user's incorrect interpretation of data, but from confusion caused by distortions introduced by map projections and poor contrast choices. While a determined user might eventually arrive at a correct interpretation, the task is made more difficult by poorly designed maps. By carefully considering the decisions during the map-making process, maps can be made that enhance user understanding of spatial information.

Maps were the dominant medium for visualizing geographic information for the first few decades of geovisualization research, making the literature on user-centered map design abundant. That research forms the basis for most modern-day cartography courses. However, geographic visualization is not limited to studies within cartography. Since 1971, geographic visualization has expanded to include much more than modifications to map symbology. In the

1990s, map animation became popular within the field of geographic visualization, and with continued advances in technology, remains popular today. Map animations are an appealing presentation format because they allow viewers to see changes over space and time while also emphasizing location (DiBiase et al.1992; Harrower and Fabrikant 2008). Three-dimensional visualizations lighten the viewer's cognitive load by no longer requiring the user to create the 3D visualization conceptually (Wood et al. 2005). Flow maps take graphs, or networks of related nodes connected by edges, and create a way of visualizing geographic data that emphasizes relationships between places instead of the places themselves (Koylu and Guo 2017; Rodgers 2005).

The Web is also responsible for the production of new geographic visualizations. Web-GIS has allowed for the development of interactive maps displaying large amounts of geographic data to be accessed by people all over the globe. As Kraak (2004) points out, this has changed the role of the map. In the context of the Web, the map medium is no longer just a means of seeing spatial patterns. Kraak shows how the map can now be both a search engine and an interface connecting users to other geographic and non-geographic information. Robert Roth has contributed substantially to geographic visualization research focused on interactive maps (Roth 2012, 2013a, 2013b, 2015). From developing a taxonomy of primitives for interactive cartography to user-centered design techniques for interactive mapping, Roth has begun developing similar cartographic theories that geographers use for static maps, but for interactive maps instead.

The Web has also led to the development and use of geovisual analytics tools and spatial mashup technologies. Geovisual analytics refers to the use of interactive interfaces to perform analytical reasoning on spatial information (Robinson 2017). These tools have a variety of applications but focus on geocomputation in order to better analyze and understand spatial information (Roth, Ross, and MacEachren 2015). Spatial mashup technologies pull spatial data from multiple sources and funnel these data to a single Web application where the data are then represented through maps. This form of geovisualization is common in disaster management (Karnatak et al. 2012).

Even with just these few examples of new ways of viewing geographic information, it is clear that substantial progress has been made in geographic visualization research. However, while there has been progress, many of the mediums for viewing geographic data developed in the past five decades are heavily map-centric. These visualizations have added to our knowledge of user-centered design and, arguably, improved our ability to interpret data to better understand spatial data and processes. However, with the notable exceptions of Anselin (2000) and Anselin et al. (2004; 2006), there are relatively few examples of research exploring the benefits of multiple representations of geographic information to explore geographic problems. Additionally, within the realm of spatial problem solving, no examples of research can be found that evaluate how such visualizations might impact both the speed and accuracy of decisions.

The lack of evaluations of visualizations engaging with multiple representations is especially surprising considering Monmonier (1989) previously showed the value of connecting multiple representations of spatial data in the same presentation through "brushing". Software programs such as Esri's ArcGIS have long allowed users to "brush" data, or interactively select elements of spatial and aspatial data and see those choices reflected in maps, tables, and graphs. However, ArcGIS is most frequently used in educational or professional settings given the high cost of the software and the training needed for its effective use. Anselin's pioneering work with GeoDa™ is the first example of a user-friendly and non-technical program that allows for multiple representations of the same data (e.g., maps and graphs) along with results from various spatial statistics.

A digital framework for connecting multiple and diverse sources and representations of spatial and aspatial data has yet to be practically transferred to a user-friendly and non-research application. The proposed GID presented in this study attempts to serve as that collection of integrated visualizations to provide a comprehensive view of key data needed to make quick, accurate decisions about a geographic process.

2.2.3 Progress in Geographic Visualization Evaluation

The final gap in the literature that the development and evaluation of the GID will address pertains to the evaluation of geographic visualizations. Evaluation is a vital step in visualization research as it allows researchers to see if a specific presentation format is effective and what conditions or characteristics influence its degree of effectiveness. Elwood (2010) points out that, when considering the context of interactive maps on the Web, the perspective of the audience changes. Advances in technology (e.g., Adobe Photoshop) and the use of volunteered geographic information (Goodchild 2007) can make users skeptical of data represented on the Web. Visual epistemologies are changing; seeing is no longer enough to justify knowing. As geographic visualizations continue to develop in tandem with technological advances, changes in the users of, and the role played, by a given visualization should also be expected to make the continual evaluation of new geographic visualizations very important.

The effectiveness of visualizations is evaluated through usability and utility testing. Usability testing evaluates the overall effectiveness of a visualization by testing its usefulness to a user. Utility testing is often used to help refine presentation formats as it focuses on the specific functions of a visualization that a user needs to complete a given task. Both usability and utility testing employ task completion exercises which involve users completing a series of tasks that require them to engage with the presentation.

Usability testing determines whether or not a user is accurately interpreting the data represented in a presentation. North (2006) provides a review of methods used to measure insight when evaluating visualizations, and these methods can be found in geographic visualization research as well. The first of these methods is controlled experiments, using what North calls "benchmark tasks". Benchmark tasks in controlled experiments can be simple tasks that are either multiple choice or have a definitive correct short answer (Tobón 2005; Bass and Blanchard 2011), but they can also be more complex, open-ended questions that force users to articulate their answers rather than relying on multiple choice options (Rinner 2007). The drawback of controlled experiments using benchmarks tasks, especially simple tasks, is they are not always effective at measuring insight because they lead the user step-by-step through the visual presentation. For this reason, there is a third approach to visualization testing that eliminates benchmark tasks completely and simply requires users explore the data in the visualization and report their findings. A few initial, open-ended questions may be used to get the user started, but ultimately, the insights gained are determined by the user, providing the researcher with a better understanding of how the presentation format will work in a real-world application. The drawback of eliminating benchmark tasks is that results from such experiments can be difficult to synthesize and time-consuming to interpret. Controlled experiments allow for easier generalizability and for rigorous statistical analysis, while the elimination of benchmark tasks yields deeper insights.

However, regardless of the method, or combination of methods, used for evaluating a geographic visualization, the sole focus of many evaluations is to measure insight, or accuracy of data interpretation, in relation to a set of independent variables. The evaluation method for the GID proposes to not only measure insight through task response accuracy but to also measure the speed of understanding through task completion time. Using time in conjunction with accuracy as a measure of effectiveness recognizes that, in real-world applications, it is not only the accuracy of our understanding that is important, but also coming to accurate conclusions in a timely and efficient manner. While studies outside of geography have evaluated presentation formats using time and accuracy as dependent variables (Vessey 1991; Dennis and Carte 1998; Speier 2006), within geography, the only visualizations to be tested using such methods are map animations and flow maps (Koussoulakou and Kraak 1992; Cutler 1998; Griffin et al. 2006; Koylu and Guo 2017).

The evaluation of the GID used in this study will also incorporate research from spatial thinking. As geographic illiteracy is widespread in the United States (National Research Council 1997), spatial thinking abilities vary from individual to individual and being able to understand differences in performance in relation to an individual's spatial thinking ability is important. Further, research in spatial thinking can be used to understand if the GID promotes spatial thinking. As poor geographic literacy is an obstacle educators have been seeking to overcome, there is much to be gained from developing geovisualizations that facilitate spatial thinking at more advanced levels.

2.3 Method for Evaluating the GID

The following section provides a specific method for evaluating the GID and gives a detailed explanation of how the method was established. By identifying four broad areas of research to form the basis of the GID, literature and subfields to draw upon for an evaluation method were identified (Figure 6). The four areas of research are as follows: visualization, information dashboards, geography, and education. Then, more specific areas of research were selected, as seen in the outer ring of the figure. The remainder of this section will individually address how research associated with each area can be tied together to create a robust method for evaluating the GID.

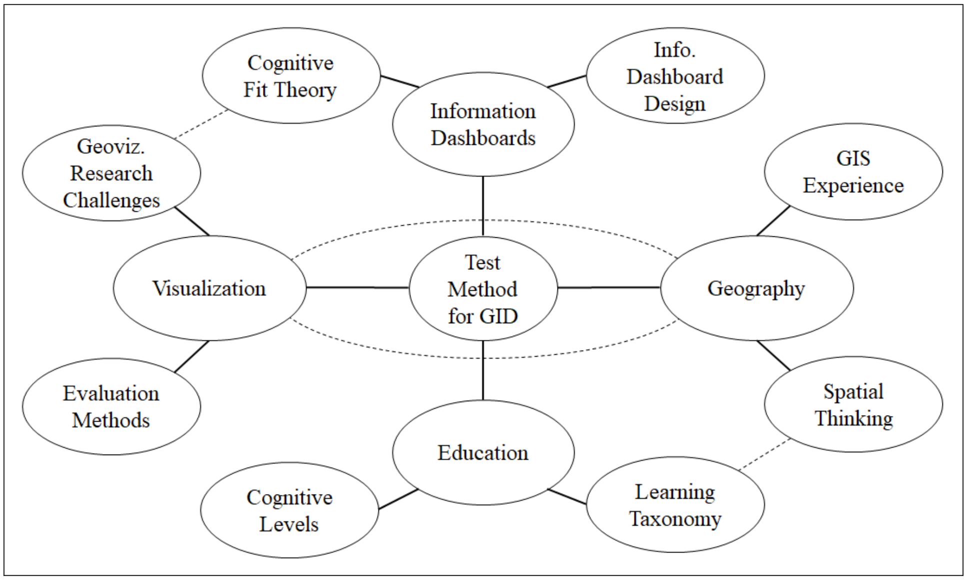


Figure 6 Concept map displaying how a test method for the GID can be developed by pulling from relevant research; dotted lines represent connections between separate fields of study.

2.3.1 Visualization

The first step in constructing a method of evaluation for a new visualization is to consider the specific research challenges associated with visualization research. MacEachren and Kraak (2001) outline core themes in geovisualization research: representation, visualization-computation integration, interfaces, and cognitive/usability issues. MacEachren and Kraak also discuss the challenges associated with each theme to aid and guide researchers seeking to make progress in the field of geographic visualization. As the focus of this section is constructing a testing method, only the latter two themes will be discussed.

The third theme, interfaces, undertakes the tasks of developing user-centered design techniques and creating interfaces that support advanced visualizations. As the GID will be an advanced, interactive visualization, this challenge is particularly important. To better understand whether or not a presentation makes sense functionally and is intuitive for users, many geovisualization studies use a two-step evaluation process. The first test evaluates the visualization's functionality through utility testing. Then, the presentation format is refined based on the feedback from the results of the first test and the second test is executed, evaluating the effectiveness of the visualization for information interpretation (Tobón, 2005). This process allows for the bugs in the interface to be worked out prior to the testing of the GID, preventing issues with the interface from inhibiting optimal information interpretation. While this theme focuses on utility testing of a visualization rather than overall effectiveness, it is still important to understand because as research on GIDs progresses, utility testing of the GID interface will become essential.

The fourth theme, cognitive/usability issues, is charged with arguably the most difficult but important of research challenges. Not only does this theme face the challenge of developing cognitive theory, but also seeks to establish methods for assessing geographic visualizations and determining the criteria under which a visualization can be considered effective or successful. To build the foundation for the GID's evaluation method and begin to confront this research challenge, studies evaluating geographic visualizations can be used as resources for methodological frameworks.

While there are many examples of evaluation methods within geovisualization (Tobón 2005; Rinner 2007; Koylu and Guo 2017), this section will focus on a book chapter by Harrower and Fabrikant (2008), as it discusses the basic method for evaluating a map animation visualization. Harrower and Fabrikant discuss how controlled experiments are commonly used to evaluate the effectiveness of map animations. Experimental testing on map animations requires the comparison of a map animation to a visualization displaying the same information in an alternate format. The alternate format should be the same in every way to the visualization being tested, except for the key characteristic that is unique to that presentation format. In the case of the map animation, it is the animation itself. Thus, researchers testing map animations use static, small-multiple maps as the comparison (Koussoulakou and Kraak, 1992; Cutler, 1998; Griffin et al., 2006). Participants in the experiment are then exposed to the map animation or the small-multiple maps and asked a series of questions to gauge their understanding. Both task response time and accuracy have been used to measure effectiveness when testing map animations. However, results of these studies have proved inconclusive.

This literature, although unable to determine the effectiveness of map animations, shows how controlled experiments can be used to evaluate geographic visualizations, and unlike most evaluation methods found in geovisualization, incorporates both time and accuracy as dependent variables. This method provides the foundation for the evaluation method of the GID.

2.3.2 Information Dashboards

As shown previously, there is very little peer-reviewed literature on the evaluation of information dashboards. When seeking guidance on how to test the effectiveness of an information dashboard, the only subject area with much experience is medicine. A few researchers have conducted experiments evaluating medical dashboard visualizations (Dolan et al., 2013; Kavanaugh et al., 2015; Dunn et al., 2016). However, given the significant difference in the nature and application of medical dashboards versus those in other disciplines, the methods used in these studies are not easily transferred to a GID.

Recognizing the lack of scholarly literature on information dashboards, Yigitbasioglu and Velcu (2012) wrote a literature review pulling together knowledge from Cognitive Fit Theory and Information Systems research to propose design and evaluation techniques for information dashboards. Cognitive Fit Theory explains the relationship between tabular and graphic representations and different types of problem-solving tasks through experimental testing (Vessey, 1991; Dennis and Carte, 1998; Speier, 2006). In the initial experiments used to test Cognitive Fit Theory, half of all participants were exposed to the tabular representation of data while the other half were exposed to the graphical representation. All participants were given the same tasks to complete, and the tasks were categorized as either symbolic or spatial. Each task a participant completes is timed and then graded for accuracy. This provides additional support for incorporating time as a variable in visualization testing, but more importantly, this method of evaluation shows how task categorization can be used to better understand the effectiveness of a presentation. By categorizing the tasks used in the experiment in some way, the researcher can observe if inconsistent results in user performance stem from an ineffective presentation format or from differences in the type of task being asked of the user. By adding knowledge obtained through Cognitive Fit Theory research to what we know about evaluation methods within visualization research, the foundation for evaluating the GID is further constructed.

2.3.3 Geography

Spatial thinking has long been an essential part of geography as a discipline. From Pattinson's (1964) spatial tradition to Murphy's (2014) cross-cutting spatial relationships theme, spatial thinking has been, and is, a pillar of geography. Golledge (2002) even went so far as to refer to spatial thinking as the very nature of geographic knowledge. Geographers are not the only ones to recognize the importance and uniqueness of spatial thinking. In 1983, Howard Gardner, a psychologist and professor at Harvard, published *Frames of Mind* which explained his Theory of Multiple Intelligences. This theory is based on the premise that there is not a single form of intelligence, but instead different ways of being intelligent. Gardner initially identified seven intelligences, one of which was spatial intelligence.

Spatial thinking and spatial memory, though closely related and in some contexts used interchangeably, are differentiated in this thesis. Spatial memory refers to memory for spatial information, such as the route from home to school or the placement of specific cards while playing memory match. Spatial thinking refers to the cognitive skills needed to comprehend spatial information, use spatial representations, and conduct spatial reasoning (National Research Council 2006). While engaging in spatial thinking requires the use of spatial memory, spatial thinking also infers the use of a broader set of cognitive skills needed to understand concepts of space and the relationships between places.

The geographer who has contributed the most to the knowledge of spatial thinking skills and concepts is undoubtedly Reginald G. Golledge. Golledge published a wide variety of papers about spatial thinking ranging from a study on people's understanding of spatial thinking concepts to a hierarchal categorization of spatial thinking concepts for the development of a spatial ontology (Golledge 1992, 1995, 2002; Golledge, March, and Battersby 2008). Others have also added to the conversation on the nature of spatial thinking. Phil and Carol Gersmehl (2007) have completed research that supports the notion that spatial thinking can be developed as a skill over time with proper education.

This has resulted in many studies that have addressed how geography education can be used to develop spatial thinking skills (Jo and Bednarz 2009; Jo, Bednarz, and Metoyer 2010). Jo and Bednarz developed a spatial thinking taxonomy to better understand the degree to which geography textbook questions facilitate spatial thinking (Jo and Bednarz 2009). Lastly, Lee and Bednarz (2012) developed a Spatial Thinking Ability Test (STAT) as an example of one way to evaluate an individual's spatial thinking abilities.

When considering a visualization that displays geographic information and is aimed at addressing a geographic problem, it is necessary to consider how an individual's spatial thinking ability may impact one's performance when using the visualization. In order to evaluate participants' spatial thinking abilities, the STAT can be used as a gauge of participants' spatial thinking skills, whether those skills are inherent or learned. The STAT is composed of sixteen multiple choice questions designed to test different aspects of spatial thinking as defined by the Committee on the Support for Thinking Spatially (2006), Gersmehl (2005), and Golledge (2002). Because the STAT tests abilities ranging from comprehending orientation and directions to overlaying and dissolving maps, regardless of the specific geographic application of the GID, the STAT is likely to evaluate the concepts needed to analyze the geographic information presented. Incorporating STAT as a measure of spatial thinking ability into the evaluation method for the GID will provide insight into whether the GID is of benefit to certain groups of spatial thinkers.

In addition to a participant's spatial thinking ability, a participant's GIS experience may also prove influential when evaluating a geographic visualization, especially the GID. Studies have shown that participation in GIS courses leads to improved spatial thinking abilities (Kim and Bednarz 2013; Lee and Bednarz 2009), but it has also been established that even non-GIS geography courses that incorporate GIS into the classroom can lead to improved spatial thinking abilities (Jo et al. 2016). In both cases, the degree of improvement is influenced by the students' performance in the course.

An understanding of a user's GIS experience is also necessary because the GID is an interactive, computer-based application that displays geographic data. Users who have experience with GIS are likely to be more comfortable with the interface and arguably more confident using it. Regardless of a user's performance in a GIS course, this experience is likely to prove advantageous. By collecting data on both spatial thinking ability and GIS experience, the method of evaluation for the GID can better determine the effectiveness of the presentation by accounting for variables that may influence user performance in both accuracy and speed.

2.3.4 Education

The final research area needed to complete the evaluation method for the GID is education. This section ties together ideas from Cognitive Fit Theory, spatial thinking, and education to show how tasks can be classified to better understand variations in performance.

Visualizations are used to communicate data to a user, who then interprets data and learns from the processed information. Different visualizations require users to engage with different learning processes depending on the function or goals of the particular presentation format. The GID is specifically designed as a comprehensive visualization, compiling data from multiple sources for the purpose of aiding in decision-making for spatial problems. This calls for the user to see the data, make sense of it, then come to a conclusion or decision. By considering different cognitive processes, we can better understand the degree to which a GID aids or impedes a user's performance at different stages in the thinking process. However, as the data the users are required to comprehend is geographic, aspects of spatial thinking need to be incorporated into our understanding of cognitive processes.

While there are many examples of learning taxonomies that distinguish different cognitive levels (Bloom, 1956; Moseley et al., 2005; Stahl and Murphy, 1981), these taxonomies consider learning as a general concept and do not specifically address *spatial* learning. As mentioned in a previous section, Cognitive Fit Theory shows the benefits of categorizing tasks used in a controlled experiment, but applying the same task classification scheme as Cognitive Fit Theory is not optimal as it is not specifically suited for presentations of geographic data. Instead, the spatial thinking taxonomy developed by Jo and Bednarz (2009) can be utilized for categorizing the tasks according to the level of spatial thinking they facilitate. This not only allows for tasks to be classified using cognitive levels but also considers the influence of spatial learning.

The spatial thinking taxonomy consists of three primary categories which were adapted from the elements the National Research Council (2006) identified as comprising spatial thinking: spatial concepts, representation tools, and reasoning processes. The spatial concepts are categorized as non-spatial, primitives, simple-spatial, and complex-spatial, a classification first proposed by Golledge (2002). The primitive concepts are place identity, location, and magnitude. These concepts are considered to be the basic concepts needed for an existence in space. Simple-spatial concepts, such as distance and adjacency, are then a combination of primitive concepts, and complex-spatial concepts, such as distribution and density, are combinations of simple or primitive concepts.

The second primary category of the taxonomy is tools of representation, which are partitioned into two groups: use and non-use. Use means that to answer a question, the student must use a visual representation such as a map or graph, while non-use means the student must use a verbal representation.

The final primary category involves cognitive processes. The three levels of cognitive processes comprising this category come from Costa (2001) but also align with the functions of spatial thinking proposed by the National Research Council (2006). Costa identified three levels of thinking: input, processing, and output. Input level thinking simply requires the student to recall or gather data. The processing level of thinking requires the student to analyze and classify input data, and output level thinking necessitates the creation of new knowledge. At this level the student must evaluate the connections made during the processing level and come to a generalization about the synthesized information. Through the use of these three primary categories, tasks can be classified according to the level of difficulty, with lower-order tasks engaging minimally with spatial thinking and higher-order tasks requiring highly developed spatial thinking skills.

2.3.5 A Method for Evaluation

Through the combination of research from visualization, information dashboards, geography, and education, a basic framework for evaluating the GID is established. Visualization provides the foundation for controlled experiments measuring both time and accuracy as dependent variables, while information dashboard research draws upon Cognitive Fit Theory to contribute methods for task categorization. Geography and education connect ideas on spatial thinking and cognitive process that are imperative to include in the evaluation as they allow other variables that may influence performance to be incorporated into the experimental design.

2.4 Conclusion

This chapter proposes the development and evaluation of a new geographic visualization, a geographic information dashboard. The development and evaluation of the GID will address several gaps in peer-reviewed literature across multiple disciplines. The literature resulting from GID research will also contribute to our understanding of information dashboards by introducing a practical and geographically-oriented visualization tool within the larger body of geovisualization research. Lastly, the evaluation method for the GID will add to our knowledge on how aspects spatial thinking influence performance when using a geographic visualization.

This chapter, secondly, provides a framework for how to evaluate the GID, so researchers who desire to pursue research of the GID can show its value through testing its effectiveness. By way of connecting research from visualization, information dashboards, geography, and education an evaluation method was established. Altogether this chapter illustrates why the scientific community needs a GID and how we can use scholarly research to build a rigorous method for evaluating it.

Chapter 3 - Methods

3.1 Introduction

This chapter describes the methods used to compares a new, integrative geographic visualization, the geographic information dashboard (GID), to a non-integrated, tab-based visualization to determine the effectiveness of the GID. The two presentation formats displayed the same geographic information, but in different ways. After the development of the GID and alternative visualization, Kansas State University students were either exposed to the GID or the tab-based format and provided a series of tasks that gauged their understanding of the information. As information dashboards, in general, are believed to lead to quick, accurate decision-making, the evaluation of performance while using the two different presentation formats was based on speed and accuracy of task response. Additionally, while some evaluations of geographic visualizations measure effectiveness through speed and accuracy of interpretation, this study also considers how spatial thinking abilities and task type influence performance when using the GID. The research questions addressed by this study are as follows:

1. What factors are most influential in determining participant performance when using the visualizations?
2. Assuming equal spatial thinking abilities, does the geographic information dashboard improve the performance-related metrics for difference spatial thinking tasks?
3. Does the GID allow participants to perform better than their spatial thinking ability would suggest?

To address the research questions posed in this study, a controlled experiment using a two-by-three-by-three factorial design was employed (Table 1). Experiment participants were either exposed to the GID or a tab-based visualization, then given a series of tasks designed to gauge understanding of the data presented. Differences in presentation and spatial thinking tasks and abilities were then used to analyze performance through both task completion time and response accuracy when using the two presentations.

Table 1 Factors and sub-groups for 2 x 3 x 3 factorial design.

Factor	Sub-groups
Presentation (2)	GID
	Tab-based
Spatial thinking ability (3)	Good
	Average
	Poor
Task type (3)	Input
	Process
	Output

Exploratory and statistical analyses were conducted, with both treating the dependent variables, task completion time and response accuracy as separate variables and using the Inverse Efficiency Score (IES) (Bruyer and Brysbaert 2011) to combine the two dependent variables into a single variable representing overall performance. Time and accuracy were first analyzed separately to allow for differences in the two measures of performance to be explored individually. However, as task completion time and response accuracy are inextricably intertwined (e.g. increased task accuracy is typically achieved by an increased amount of time spent on a task) and as the relationship between the two is variable from individual to individual (e.g. fast and slow test-takers, students with test anxiety), the IES was used to quantify the two variables into a single measure representing overall performance.

Exploratory analyses were conducted using boxplots and interaction plots to provide visual conceptualizations of the data. Statistical analyses were then executed in three steps. First, 4 one-way ANOVAs were completed using the IES data to determine which independent variable was most influential on performance. Second, a difference of means test was used to establish if there were significant differences in performance on output tasks between the two presentations. Lastly, the Kolmogorov-Smirnov (KS) test was used to compare performance metric distributions between participants using different presentations and belonging to different spatial thinking groups.

3.3.1 Independent Variables

Three independent variables were included in the analysis: type of information presentation, assessed spatial thinking ability as determined by performance on the Spatial Thinking Abilities Test (Lee and Bednarz 2012), and task type.

The first independent variable was the type of information presentation, either the dashboard presentation or the tab-based presentation. Both presentations were created using a package in R statistical software called R Shiny and were published as Web applications. The application for the visualizations was land management at the Fort Riley Military Reservation in Fort Riley, Kansas. As land management is a multi-faceted problem within geography, it was deemed an appropriate application to test the dashboard visualization, as it is a presentation that integrates data from multiple sources. The more specific focus of the land management application was to analyze grassland vegetative development at Fort Riley, Kansas, during the year 2012.

The dashboard presented a total of six widgets on a single screen (Figure 7). In this context, a widget refers to a graphic within a Web application that can be manipulated and allows the user to access data. The dashboard consisted of two interactive map widgets created using Leaflet, an open-source JavaScript library for interactive maps. The interactive maps displayed Normalized Difference Vegetation Index (NDVI) data derived from 16-day maximum value NDVI composite images collected by the MODIS sensor onboard the Terra satellite (MOD13Q1). There were also two histogram widgets that represented the distribution of NDVI values presented in each of the respective maps. Lastly, there were two line-graph widgets displaying phenology and weather. The six widgets were controlled by three drop-down menus. For each map widget, there was a drop-down menu that allowed the user to select data for a specific time of the year and have it displayed on the map, while also automatically updating the histogram associated with each map to display the aspatial distribution of the selected mapped data. The third drop down menu allowed the user to select a time of year and see the pattern of vegetation phenology (i.e., a line graph of average NDVI) and weather (i.e., combined line and bar graphs representing temperature and precipitation) from day of year (DOY) 1 through the user-selected date.

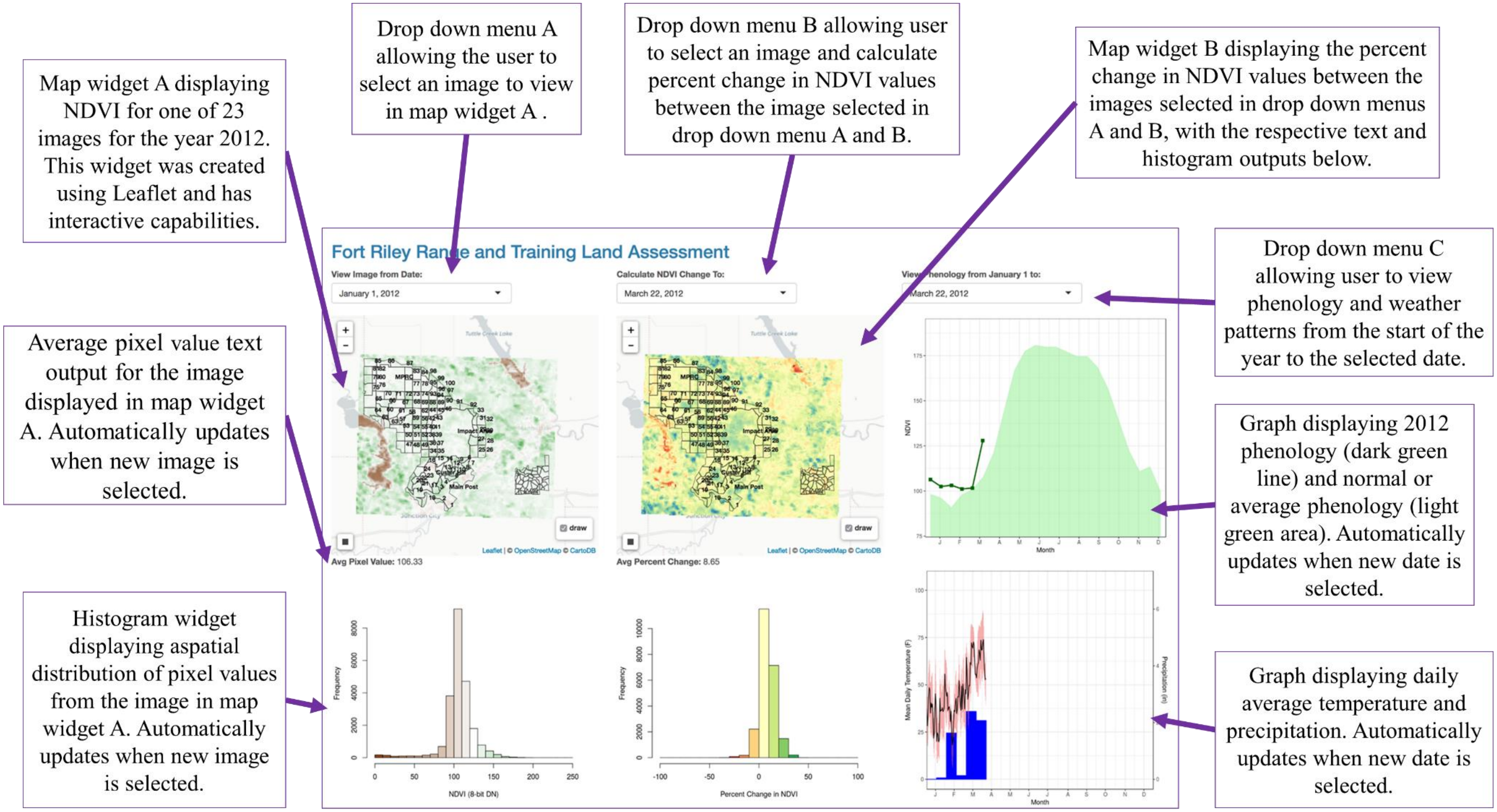


Figure 7 The GID Web application used in the experiment.

The tab-based presentation was non-integrative and displayed each of the six widgets on separate tabs of the application (Figure 8). Since each widget was displayed on its own tab, they were all controlled by their own drop-down menus. Developing the tab-based visualization in this manner eliminated the side-by-side comparisons and automatic updating of the dashboard presentation. With the exception of the added drop-down menus for the tab-based format, both presentations had the same functions and displayed the same data.

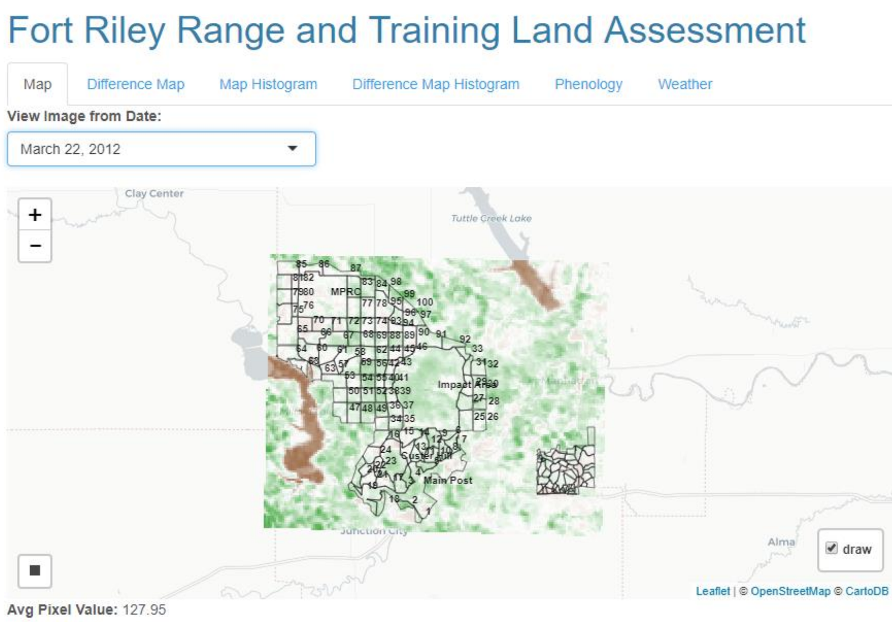


Figure 8 Tab-based Web application used in the experiment.

The second independent variable was spatial thinking ability. Assessing spatial thinking ability was determined through a combination of a participant's spatial thinking ability and his/her GIS experience. Spatial thinking ability was necessary to understand a participant's ability to think about and solve spatial problems. GIS experience was used as an indicator of exposure to geographic data and experience using geospatial technologies.

A modification of the Spatial Thinking Abilities Test (STAT) developed by Lee and Bednarz (2012) was administered prior to a user's exposure to the GID to determine a participant's spatial thinking ability. The participant's percentage correct out of sixteen questions was then recorded as their STAT score.

Experience using geographic data and geospatial technologies was determined through the collection of data on participants' previous or current course enrollment. If participants had completed two or more GIS courses or were currently enrolled in their second GIS course, they were considered to have sufficient experience to prove advantageous when using the geographic visualization.

The final independent variable was the type of task. Tasks were categorized based on the type of cognitive process required to complete a given task. The cognitive processes were taken from Costa (2001) and consist of three levels: input, processing, and output.

Type of task was used as a proxy for evaluating facilitation of spatial thinking. The spatial thinking taxonomy developed by Jo and Bednarz (2009) can be used to classify tasks ranging from simple and complex spatial thinking tasks using three primary categories. However, as the tasks for the visualizations were formulated and refined, it became clear that, for this specific study, the type of cognitive process involved in responding to the task was the only necessary category from the taxonomy to apply to the tasks. Firstly, as all tasks involved the use of a visual medium, the use or non-use of a representation became irrelevant. Secondly, the specific tasks asked in this study increased in complexity of spatial thinking concepts as they increased in complexity of the cognitive process. This means that input level tasks involved simple spatial concepts and output level tasks involved complex spatial concepts. For this reason, it was determined that type of task was sufficient for measuring the facilitation of spatial thinking.

An online survey comprised of twenty total tasks was created using Qualtrics. The survey was comprised of four input level tasks, eight processing level tasks, and eight output level tasks. Not as many input level tasks were asked because the more complex tasks were of greater interest, while the simpler, input level tasks were only needed to ensure that participants did not provide incorrect answers for more difficult questions because of inaccuracies stemming from simpler questions. Survey tasks were informally piloted on a second researcher, a geography undergraduate, and a non-geography professional to ensure task clarity and relevance.

3.3.2 Dependent Variables

Two dependent variables were selected for analysis: task completion time and task response accuracy. The time it took a participant to answer a given task was recorded in seconds as part of the Qualtrics survey. Each question was displayed on its own page; time began when the participant opened the page and ended when the participant clicked on the next question, taking them to the next task on the following page. The timing component was not visible to the participant.

Response accuracy was determined by awarding points to each task based on its complexity. If tasks could only have a clear right or wrong answer, they were awarded one point, whereas tasks of greater complexity were awarded two points. This allowed for participants who provided answers to more difficult tasks that were correct but did not address the entire task, to be awarded partial credit. The researcher graded all responses twice. Additionally, two participant surveys from each presentation group, one scoring low in accuracy and one high in accuracy, were provided to a second researcher to be independently rated to ensure consistent enforcement of the grading rubric.

3.3.3 Experiment Procedures

Participants were recruited from undergraduate geography courses and included a mixture of undergraduate and graduate students from a variety of academic majors (IRB proposal number 9006). While all participants were enrolled in at least one geography course, all had different levels of experience with spatial thinking and geospatial technology software. To incentivize students to enroll in the experiment, instructors of the participating geography courses offered extra credit to participants.

Following the resource equation for factorial design experiments from Mead (1988), the total number of participants minus the number of treatments should be between 10 and 20. It was determined that a minimum of thirty-nine participants were required for this study:

$$E = N - B - T$$

where:
E = error degrees of freedom
N = total degrees of freedom
B = blocks degrees of freedom
T = number of treatments

Since the experiment conducted was a non-blocked design, the variable B is eliminated. Assuming an optimal value for error degrees of freedom is 20 and using 18 as the total number of treatments (2 x 3 x 3 factorial design), the total degrees of freedom is 38. Since total degrees freedom is equal to sample size minus one, a minimum of 39 participants was required.

A total of forty-one participants enrolled in the study (Table 2). Twenty-two participants used the dashboard format, while 19 participants used the tab-based format. Five experimental sessions were conducted, with each session comprised of between 4 and 12 participants. The dashboard format was used for three sessions, while the tab-based format was used for two sessions. The same researcher conducted all experimental sessions and followed a standard script.

Table 2 Descriptive data for participant group.

Participant Characteristics		
Gender	Male	19
	Female	22
Educational Level	Undergraduate	38
	Graduate	3
Major	Engineering	6
	Agronomy	4
	Anthropology	4
	Education	4
	Geography	4
	Geology	2
	Biology	2
	Economics	2
	Apparel and Textiles Marketing	2
	Business	2
	Other	9
GIS Experience	0 Courses	19
	1 Course	12
	2 Courses	10

The experiment began with the STAT (administered on paper), and participants were given an unlimited amount of time to complete the test. After all participants completed the STAT, the participants listened to a five-minute presentation on the data that would be presented in the experiment and a description of the graphs that would comprise the dashboard and tabbed presentations. The participants were also provided a handout containing the content from the presentation for note-taking and reference throughout the experiment. This was a vital step in the experiment as it ensured all participants, regardless of background, had a working knowledge of the data used in the visualization and why it was important.

As shown previously by Jarvenpaa and Dickson (1988), there is a need for training and practice before the measurement of performance. Therefore, following the presentation, a demonstration was conducted to illustrate how to use all the features in the visualization. After the demonstration, participants were given five minutes to navigate the visualization on their own and ask any questions about the functionality of the Web application.

After the overall introduction to the problem and training, participants were given an unlimited amount of time to complete an online survey with twenty open-ended questions to answer using the visualization . The experiment was completed once the online survey was submitted.

In addition to the extra credit received for participation, participants had the option to compete for a performance-based reward of $10 for providing the most accurate responses out of all participants in their experimental session. Performance-based rewards have been shown to increase participant effort and accuracy and were used to encourage participants to strive for accuracy in their responses, rather than solely focusing on providing quick responses (Creyer et al. 1990).

3.3.4 Analysis Procedures

The first step in the analysis process was to identify outliers in the data set. Using an adjusted quantile plot and Mahalanobis distances, 2 of the original 41 participants were eliminated. After outliers were eliminated, data on total task completion time and total accuracy for all 20 questions were normalized to a scale of 0 to 1 range to eliminate units of measurement and allow for easier comparisons between and within variables. The total time spent answering questions from a specific cognitive level (input, processing, output) was also normalized, along with the associated percentage of correct responses for each cognitive level.

To determine a participant's spatial thinking ability, STAT scores were normalized then separated into three categories of spatial thinkers: poor, average, and good. Classification for the STAT scores was determined using a quantile classification with three breaks. Once participants were labeled as "poor", "average", or "good" spatial thinkers, those reporting taking two or more GIS courses were moved from their current class to the one immediately above it. A total of 10 participants reported taking a minimum of 2 GIS courses. This resulted in three participants from the poor class being moved to the average class, while three participants from the average class were moved to the good class. The remaining four participants were already in the good class, and as such could not be moved to any higher class. This method of determining participant spatial thinking ability resulted in a total of 15 participants classified with poor, 9 with average, and 15 with good spatial thinking ability.

Once the data were organized, an exploratory analysis was conducted on the normalized data using box plots and interaction plots to visualize relationships within the data. These graphics displayed the two dependent variables separately and as a single dependent variable using the Inverse Efficiency Score (IES) (Bruyer and Brysbaert 2011) to combine time and accuracy. IES is calculated using the following equation:

$$IES = \frac{Response\ Time}{Percent\ Correct}$$

Using the IES, low data values represent participants achieving high accuracy in less time, while high data values represent participants achieving low accuracy in more time. The IES data was then used further exploratory analyses, and in the statistical analysis using multiple one-way ANOVAs. This was followed by a difference of means test and the Kolmogorov-Smirnov (KS) test.

Chapter 4 - Results

Exploratory analyses, treating task completion time and accuracy separately, showed that in general, participants using the dashboard presentation, as opposed to the tab-based presentation, completed tasks more quickly (Figure 9) and with greater accuracy (Figure 10). Additionally, there were notable differences in both speed and accuracy, regardless of presentation, when comparing participants of different spatial thinking ability classes. While differences were less distinct for time, accuracy decreased as spatial thinking ability decreased.

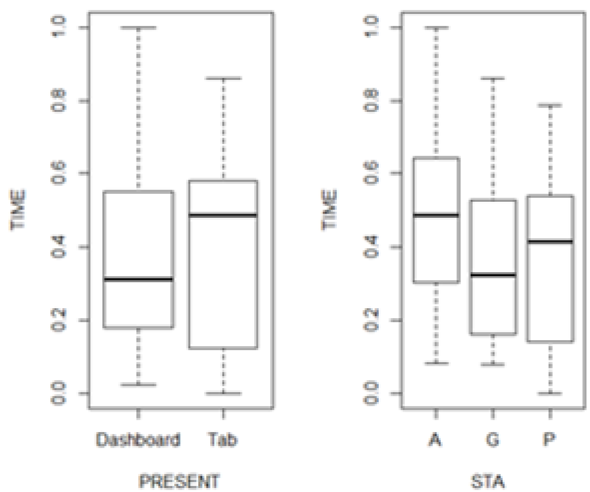


Figure 9 Differences in time between the GID (Dashboard) and tabbed (Tab) presentation (left) and between different spatial thinking ability groups (A (average), G (good), P (poor)) (right); all time data are normalized.

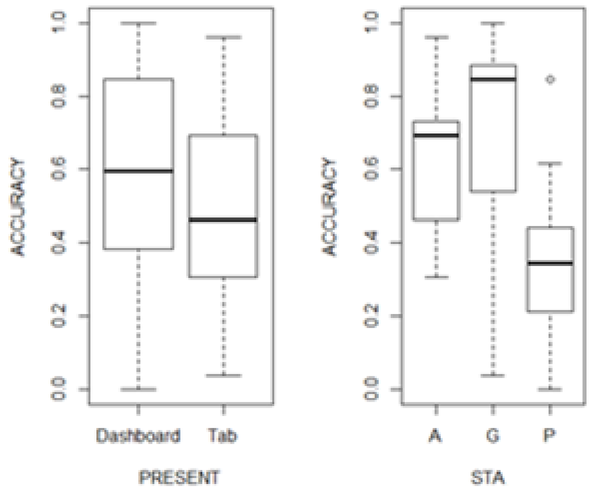


Figure 10 Differences in accuracy between the GID (Dashboard) and tabbed (Tab) presentation (left) and between different spatial thinking ability groups (A (average), G (good), P (poor)) (right); all accuracy data are normalized.

Further analysis using interaction plots showed that when comparing presentation type and spatial thinking ability in relation to time, there was only a small decrease in task completion time for participants using the dashboard. The most notable difference in task completion time being for good spatial thinkers (Figure 11). Comparing the same two independent variables in relation to accuracy, there was only a slight difference in accuracy between presentations for poor (slightly decline in performance) and average (slight improvement in performance) spatial thinkers, but for good spatial thinkers there was an increase in accuracy of >20% when using the dashboard presentation.

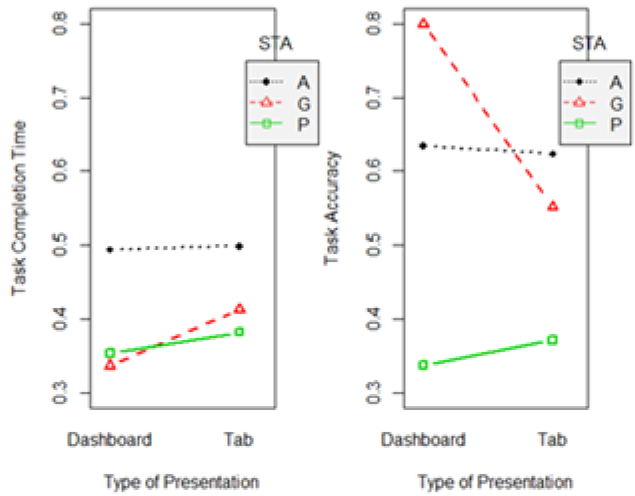


Figure 11 Interaction plots displaying changes in time (left) and accuracy (right) between the GID (Dashboard) and tabbed (Tab) presentation for each spatial thinking ability group (A (average), G (good), P (poor)) for all tasks; all time and accuracy data are normalized.

To visualize how time and accuracy varied in relation to all independent variables used in this study, including task type, separate interaction plots were created for input, processing, and output tasks. For input tasks, there was an obvious decrease in task completion time for all spatial thinking classes when using the dashboard presentation (Figure 12). However, while there was a notable increase in accuracy for good and average spatial thinkers when using the dashboard, there was a surprising decrease in accuracy for poor spatial thinkers using the dashboard.

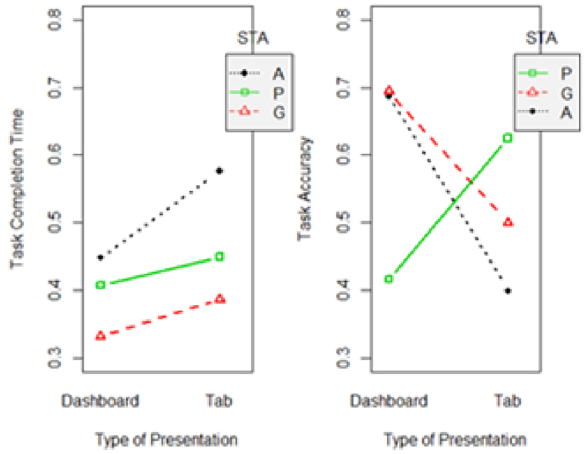


Figure 12 Differences in time (left) and accuracy (right) between the GID (Dashboard) and tabbed (Tab) presentation for each spatial thinking ability group(A (average), G (good), P (poor)) for input tasks; all time and accuracy data are normalized.

For processing tasks, we observed a decrease in time when using the dashboard presentation for poor and good spatial thinkers, but a slight increase in time for average spatial thinkers (Figure 13). Then, while good spatial thinkers experienced a strong increase in accuracy when using the dashboard, accuracies for average and poor spatial thinkers decreased.

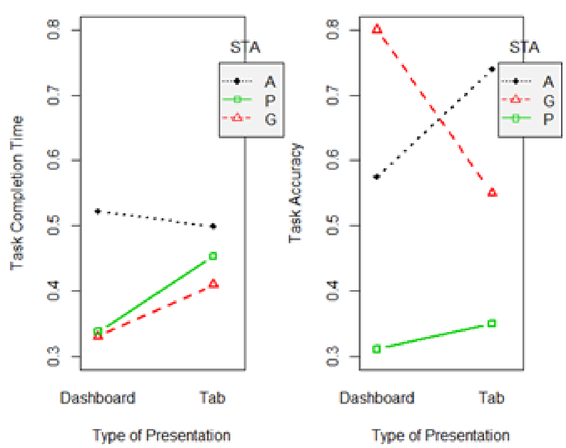


Figure 13 Differences in time (left) and accuracy (right) between the GID (Dashboard) and tabbed (Tab) presentation for each spatial thinking ability group (A (average), G (good), P (poor)) for process tasks; all time and accuracy data are normalized.

Lastly, while examining the interaction plots for output tasks, we observed a slight increase in task completion time when using the dashboard for all spatial thinking classes (Figure 14). That slight increase in task completion time, however, also resulted in improved accuracies for all spatial thinking classes, especially those in the good category.

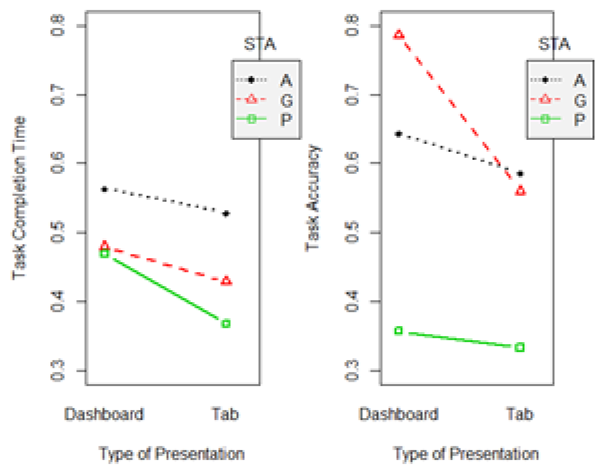


Figure 14 Differences in time (left) and accuracy (right) between the GID (Dashboard) and tabbed (Tab) presentation for each spatial thinking ability group (A (average), G (good), P (poor)) for output tasks; all time and accuracy data are normalized.

After exploratory analyses using the normalized performance data and evaluating task completion time and accuracy as separate dependent variables, further exploratory analysis was executed using the integrative IES metric, where lower values represented better overall performance. When using the IES to measure performance as a whole, there was little difference between the two presentation formats (Figure 15), though the median IES values were lower for the dashboard presentation. Additionally, the differences in performance between spatial thinking classes showed an inverse relationship between IES score and spatial thinking class.

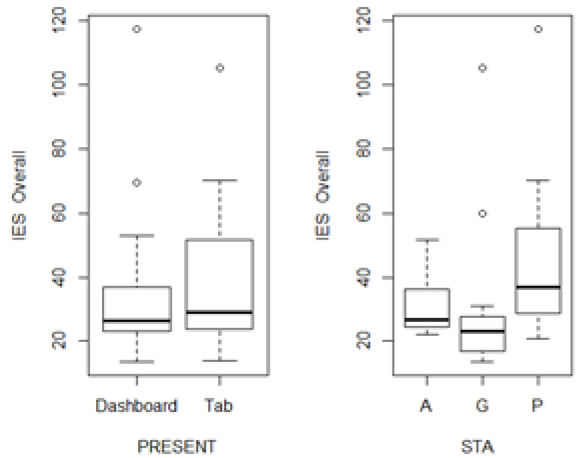


Figure 15 Differences in overall performance based on participants' Inverse Efficiency Scores (IES) between the GID (Dashboard) and tabbed (Tab) presentation (left) and spatial thinking ability groups (A (average), G (good), P (poor)) (right).

An interaction plot of the overall performance considering both presentation and spatial thinking ability indicates that good and average spatial thinkers performed better using the dashboard format, while poor spatial thinkers performed better using the tab-based format (Figure 16).

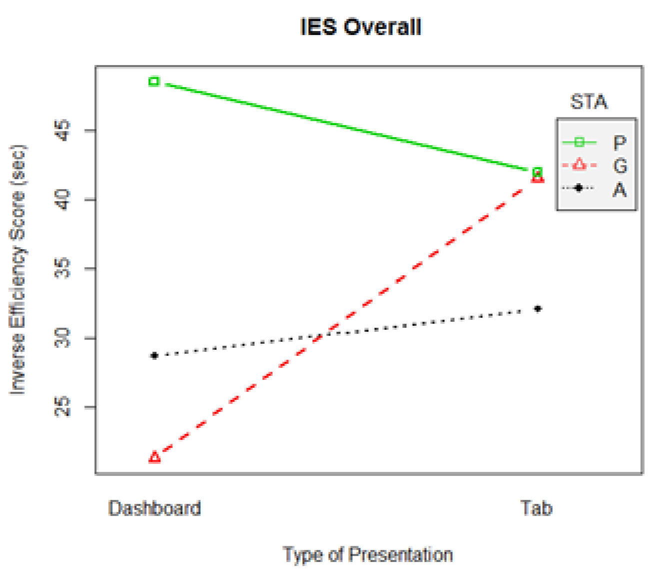


Figure 16 Differences in Inverse Efficiency Scores between the GID (Dashboard) and tabbed (Tab) presentation for each spatial thinking ability group (A (average), G (good), P (poor)) for all tasks.

Separate interaction plots were then made to compare performance between cognitive tasks. For input tasks, both good and average spatial thinkers performed better using the dashboard, while poor spatial thinkers performed better using the tab-based format (Figure 17).

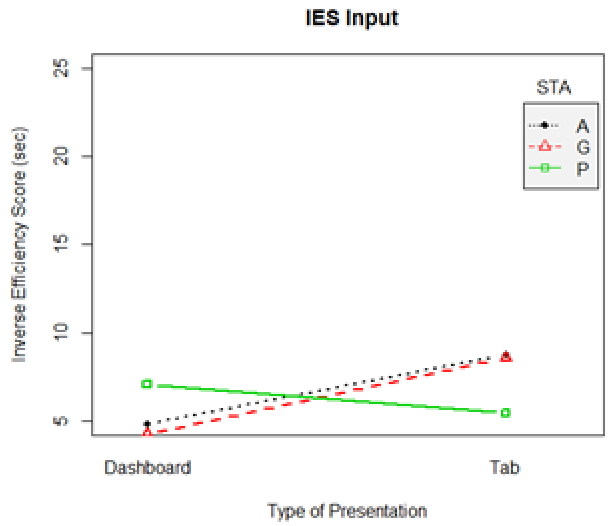


Figure 17 Differences in Inverse Efficiency Scores between the GID (Dashboard) and tabbed (Tab) presentation for each spatial thinking ability group (A (average), G (good), P (poor)) for input tasks.

For processing tasks, participants in the good and poor spatial thinking classes performed better when using the dashboard, as opposed to participants in the average spatial thinking class. The latter group performed slightly better using the tab-based presentation (Figure 18).

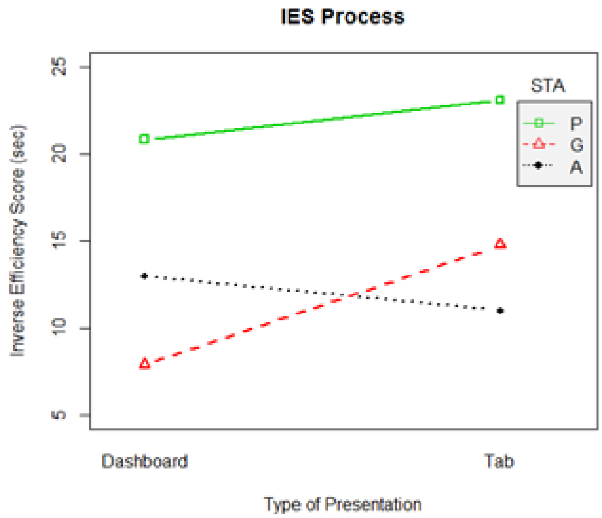


Figure 18 Differences in Inverse Efficiency Scores between the GID (Dashboard) and tabbed (Tab) presentation for each spatial thinking ability group (A (average), G (good), P (poor)) for process tasks.

Lastly, for output tasks, good and average spatial thinkers improved in performance when using the dashboard, but for poor spatial thinkers, performance improved when using the tab-based format (Figure 19).

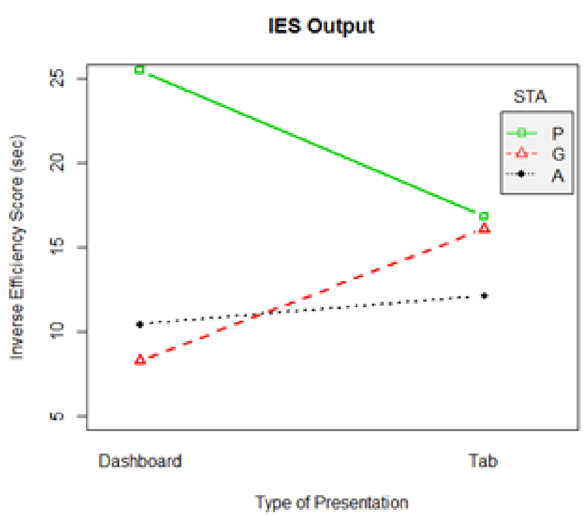


Figure 19 Differences in Inverse Efficiency Scores between the GID (Dashboard) and tabbed (Tab) presentation for each spatial thinking ability group (A (average), G (good), P (poor)) for output tasks.

Following exploratory analyses of both the normalized and the IES data, statistical analysis of the IES data was conducted. To meet the assumptions of ANOVA, the IES data were log-transformed to achieve normal distribution. Multiple one-way ANOVAs were completed to analyze performance via IES in relation to presentation, spatial thinking ability, and the interaction term between the two. The first ANOVA examined overall performance, while the following three evaluated performance by each cognitive task.

ANOVA results showed that spatial thinking ability was statistically significant for determining performance when using the visualizations, regardless of which format was used, for overall performance, and for processing and output tasks (Table 3, Table 5, Table 6). For input tasks, neither independent variable nor the interaction term between them was statistically significant, but all three were approaching significance (Table 4).

Table 3 Results of one-way ANOVA determining the influence of presentation (present), spatial thinking ability (STA) and their interaction (present:STA) on overall performance (P < 0.05).

	Df	Sum of Sq.	RSS	AIC	F Value	Pr(>F)
<none>			7.4782	-52.411		
Present	1	0.01307	7.4913	-54.343	0.0577	0.81169
STA	2	2.34832	9.8265	-45.76	5.1813	0.01104*
Present:STA	2	0.50085	7.9791	-53.883	1.1051	0.34313

Table 4 Results of one-way ANOVA determining the influence of presentation (present), spatial thinking ability (STA) and their interaction (present:STA) on input task performance (P < 0.05).

	Df	Sum of Sq.	RSS	AIC	F Value	Pr(>F)
<none>			7.6696	-51.425		
Present	1	0.62824	8.2979	-50.355	2.7031	0.1096
STA	2	1.14181	8.8115	-50.013	2.4564	0.1013
Present:STA	2	1.11554	8.7852	-50.129	2.3999	0.1064

Table 5 Results of one-way ANOVA determining the influence of presentation (present), spatial thinking ability (STA) and their interaction (present:STA) on process task performance (P < 0.05).

	Df	Sum of Sq.	RSS	AIC	F Value	Pr(>F)
<none>			10.625	-38.714		
Present	1	0.08266	10.708	-40.412	0.2567	0.61573
STA	2	3.00357	13.629	-33.004	4.6644	0.01644*
Present:STA	2	0.599	11.224	-40.575	0.9302	0.40456

Table 6 Results of one-way ANOVA determining the influence of presentation (present), spatial thinking ability (STA) and their interaction (present:STA) on output task performance (P < 0.05).

	Df	Sum of Sq.	RSS	AIC	F Value	Pr(>F)
<none>			12.96	-30.966		
Present	1	0.0258	12.986	-32.888	0.0657	0.79925
STA	2	3.3401	16.3	-26.023	4.2524	0.02274*
Present:STA	2	0.5759	13.536	-33.27	0.7332	0.48803

The second statistical test was a difference of means test. A difference in means test is used to compare performance metrics (IES, time to completion, accuracy) of users within the same spatial thinking class who used the dashboard versus tab visualization. Each performance metric was assessed for normality using the Shapiro-Wilk test, then the variances of the two datasets were assessed for equality via an F test. Depending on the results for normality and homogeneity of variances, different parametric and non-parametric tests were chosen to determine whether significant differences in the means or medians existed between the compare groups (Table 7). Regardless of test used, a one-tailed hypothesis was tested to determine whether a significantly better performance was measured for users of the dashboard visualization. Table 8 lists the p-values and difference in means computed for each of the 12 performance metrics recorded for users of the GID and tabbed visualization grouped by spatial ability class.

Table 7 Decision criteria for applying each of the four variants of a difference in means tests used to assess improvements in performance due to use of the GID.

Normal Distribution	Homogeneity of Variances	Statistical Test	Test Type	No. of Comparisons
Yes	Yes	T-Test (Equal Variance)	Parametric	22
Yes	No	T-Test (Unequal Variance)	Parametric	5
No	Yes	Wilcoxon-Mann-Whitney	Non-Parametric	5
No	No	Kolmogorv-Smirnov	Non-Parametric	4

Table 8 One-tailed test results from parametric and non-parametric methods for comparing performance metrics of dashboard vs. tab visualization users grouped by spatial thinking class. Negative values for the difference in means indicates and improvement in performance for users of the dashboard visualization. P-values in normal text = t-test (equal variance), italics = t-test(unequal variance), bold = Wilcoxon-Mann-Whitney, and bold/italic = Kolmogorov-Smirnov.

	p-values by Spatial Thinking Class		
Performance Metric	Poor (Diff Means)	Avg (Diff Means)	Good (Diff Means)
IES Overall	**0.568 (6.593)**	**0.5476 (-11.87)**	*0.1099 (-20.28)*
IES Input	0.9015 (1.640)	*0.0366 (-3.968)*	*0.0998 (-4.308)*
IES Process	0.3932 (-2.259)	0.8256 (1.973)	***0.1653 (-6.847)***
IES Output	***0.7008 (8.680)***	**0.4524 (-3.408)**	***0.1653 (-7.802)***
Total Overall Time	0.4242 (-0.0285)	*0.4913 (-0.0052)*	0.2895 (-0.0752)
Total Input Time	0.3110 (-0.0427)	0.3041 (-0.1287)	*0.3619 (-0.0545)*
Total Process Time	0.2274 (-0.1144)	0.5446 (0.0226)	0.2840 (-0.0788)
Total Output Time	0.7671 (0.1015)	0.5788 (0.0350)	0.7080 (0.0509)
Total Overall Accuracy	0.6053 (-0.0342)	0.4742 (0.0115)	0.0392 (0.2479)
Total Input Accuracy	0.9417 (-0.2083)	**0.1726 (0.2875)**	***0.4493 (0.1944)***
Total Process Accuracy	0.5917 (-0.3891)	0.7575 (-0.1650)	0.0272 (0.2500)
Total Output Accuracy	0.4238 (0.0238)	**0.2512 (0.0571)**	0.0646 (0.2262)

An examination of the calculated difference in means between users of the GID and tab visualization within each spatial ability class shows that the GID reduced task completion time or improved accuracy in 24 of 36 possible combinations (66%), though these differences approach significance (p-value <= 0.1653) in 8 instances (20%).

For participants in the poor spatial ability group, the GID appears to have the most impact on time-based metrics of performance, especially for less complex tasks. Results also suggest that, for average and good spatial abilities, users realize the most benefit from the GID through improved accuracy, especially for complex tasks. There was less observed impact on time to completion, though the GID may have helped most participants reach answers to input and processing tasks somewhat faster.

The final statistical analysis was the Kolmogorov-Smirnov (KS) test. The Kolmogorov-Smirnov (KS) test (Chakravart et al., 1967) is used to compare performance metric distributions between participants in the poor spatial ability class using the GID with those in the average class using the tabbed visualization. The KS test is based on the empirical cumulative distribution function (ECDF) of two distributions and does not depend on the tested distribution following a specific function (e.g., normality). Both a two-sided and one-sided KS test was conducted. The two-sided test was used to assess whether the distributions were drawn from the same population, while the one-sided test allowed for a more direct evaluation of whether the GID was associated with improved performance across each of the 12 metrics. For the time- and accuracy-based performance metrics, the one-sided tests considered the right (i.e., greater than) and left (i.e., less than) tails of the distribution. A Levene's test was then applied to compare the variances of each performance metric for both visualizations within the same spatial ability class (e.g., poor-dashboard vs. poor-tab) to determine whether use of the GID resulted in a significant change in the variation found within each metric (e.g., more/less dispersion around the median). The p-values for each test and comparison group are shown in Table 7.

All distributions for performance metrics evaluated are statistically similar. This suggests that GIDs may elevate the performance of users in lower spatial ability classes to a level similar to a higher class disadvantaged by using the tabbed visualization. This impact is less important for poor spatial abilities (KS test p-values are lower, or closer to a significant difference) and more pronounced for average users (KS test p-values are higher). For those in the poor class, distributions for time-based measures of performance are impacted more than accuracy (KS test p-values are higher indicating more similarity to the distribution of average-tab group). While the GID may be more effective at helping average users, the influence of the GID on time and accuracy metrics is comparable.

Table 9 Results from Kolmogorov-Smirnov and Levene's tests used to determine whether participants in a lower spatial ability class who used the GID performed similarly to those in the next highest spatial ability class who used the tabbed visualization. P-values for Levene's test is included to indicate whether there was a change in the variance associated with each performance metric for users in the same spatial ability class but using different visualizations.

	p-values			
	KS Test	Levene's	KS Test	Levene's
Performance Metric	Poor-GID: Avg-Tab	Poor-GID: Poor-Tab	Avg-GID: Good-Tab	Avg-GID: Avg-Tab
IES Overall	0.3896	0.3620	0.5524	0.4282
IES Input	0.3007	0.1350	0.5524	0.1350
IES Process	0.3806	0.4972	0.5524	0.4972
IES Output	0.1189	0.6917	0.2952	0.6917
Total Overall Time	0.2258	0.8685	0.9952	**0.0120***
Total Input Time	0.5415	0.8228	0.9238	0.8446
Total Process Time	0.4605	0.7244	0.9238	**0.0031****
Total Output Time	0.7912	0.6699	0.6952	0.6422
Total Overall Accuracy	0.1975	0.8082	0.9238	0.7370
Total Input Accuracy	0.1975	0.4346	0.9525	0.1355
Total Process Accuracy	0.1975	0.8122	0.9983	0.3885
Total Output Accuracy	0.7486	1.000	0.2365	0.4828

A statistically significant change in variance was measured for total overall time and total time to completion for processing-level tasks for the average group. Nearly significant were the difference in variance for the IES score for input tasks for both groups, as well as the total accuracy for input tasks for the average group. This hints at the dashboard visualization helping to make user responses more consistent, with the general trend in reducing the variability in responses as task complexity increases.

Chapter 5 - Discussion and Conclusion

5.1 Discussion

5.1.1 Revisiting the Research Questions

RQ1: What factors are most influential in determining participant performance when using the visualizations?

The results conclusively show that spatial thinking ability was the factor most influential in determining performance when using the visualizations. In 3 of the 4 ANOVAs conducted using the IES data, spatial thinking ability was a statistically significant factor in determining performance. Additionally, for the ANOVA conducted for input task performance, spatial thinking ability was not statistically significant, but it was very close ($p = 0.1013$). This finding is important as it calls attention to the value of including spatial thinking ability as a measure in experiments designed to evaluate geographic visualizations. Through this measure, researchers can see if a geographic presentation is best suited for specific groups of spatial thinkers and better determine an appropriate audience or user for the visualization. Future geovisualization research should also account for participant spatial abilities in assessment studies to avoid potential statistical bias in their results.

While spatial thinking ability was the only statistically significant variable influencing performance, differences in performance were observed between the dashboard and tabbed presentation formats. Though statistically significant evidence was not found in this study to support the generally held notion that information dashboards lead to faster and more accurate decisions, trends in task completion time, accuracy, and IES suggest improved performance across most spatial thinking classes for the dashboard format. Future usability tests incorporating tighter controls over the participant group or alternatives to benchmark tasks may be able to better establish a functional linkage between presentation format and performance in spatial decision-making applications.

RQ2: Assuming equal spatial thinking abilities, does the geographic information dashboard improve the performance-related metrics for difference spatial thinking tasks?

This research question was answered by interpreting the results of participant performance for output tasks. As output tasks represent the most difficult cognitive and spatial thinking tasks, by analyzing performance differences between presentation formats for output tasks, we were able to determine if the GID facilitated spatial thinking as indicated by improved performance. Results from the difference of means test show that for average and good spatial thinkers, the GID allowed for significantly better accuracy for output tasks. This provides support for the notion that the GID facilitates spatial thinking, and while the pattern is not maintained for poor spatial thinkers, these findings suggest for users with more experience in spatial reasoning, the GID format allows them to answer complex spatial thinking tasks more easily. This is important to note as it aids in determining an appropriate user for the GID and sheds light on areas of future research that can focus on developing a GID for users with less spatial reasoning experience.

RQ3: Does the GID allow participants to perform better than their spatial thinking ability suggests?

Results from the KS test reporting that all the distributions of performance metrics were statistically similar provides evidence that the GID presentation format may improve user performance beyond what is expected based on spatial thinking ability. This is an important finding as it shows that even in instances when a user was not extremely accurate, the GID still aided in the user's performance by allowing him/her to perform at the same level as a user with higher spatial thinking ability using the tab-based format. This draws attention to the GID's usefulness as a visual tool while also highlighting, once again, the value of measuring spatial thinking ability when evaluating a visualization.

5.1.2 Limitations

While this study was the first to evaluate a GID and led to many insights as to the effectiveness of a visualization designed to enhance spatial thinking and decision-making, due to constrained time and resources, there were several limitations to the work which merit further consideration. The first limitation was the sample size and participant composition. As the sample size was only 41 participants, statistical analysis could not be conducted on individual spatial thinking classes due to small sample sizes. Further, it would have benefitted this study to draw upon a more diverse participant pool. While the participants in this study had a wide variety of educational focuses, they were all within the same general age group and, likely, lacked much experience or practice in practical evidence-based problem-solving. Finally, data on participant gender was not collected or used in the analysis. In hindsight, gender may have revealed important differences in performance that, even if not significant, would have added to our knowledge base on assessments of geographic visualizations.

Further limitations include the use of benchmark tasks and lack of data collection in regards to participants' perceptions of the visualizations. This study used open-ended tasks to gauge participant understanding of the information, and while this allows for more insight to be measured than close-ended tasks, having any definitive tasks leads the participant through the information and does not allow for independent data exploration. The decision to use open-ended tasks was intentional as participants were expected to be unfamiliar with the topic and with complex decision-making, but we do recognize this choice of task as a limitation to the study. Additionally, the survey used to measure understanding did not include questions inquiring about a participant's perception of the visualization and what may have affected his/her individual performance. With the current data, we could determine the performance of every participant, but we cannot confidently say what influenced participant performance beyond the independent variables used in the experiment.

Another limitation of the study was participant apathy. Participants were motivated to attend their respective experiment sessions to receive the extra credit for participation in the study. Additionally, participants were incentivized to put in the necessary effort to provide accurate responses by giving a ten-dollar cash reward to the participant within each session who achieved the highest accuracy. However, this may not have been sufficient to motivate all participants to respond accurately to all tasks in the online survey. There is some anecdotal evidence from the experimental sessions, also revealed in the collected data, to suggest that this was a problem for many participants in the poor spatial thinking category. Though the presence of outliers were tested, and those values removed at the start of the study, additional high-quality data (data from participants who put in reasonable effort rather than providing minimal answers in order to complete the survey in less time) from participants in the poor spatial thinking class may have resulted in more outliers being eliminated from further analysis.

Lastly, as discussed earlier, there is a lack of scholarly literature published on information dashboard design. Thus, the GID suffers from a lack of theory to guide and support design decisions. While the GID was designed by two formally trained geographers with extensive knowledge of geographic data and geographic visualization theory, we recognize that another limitation of this study is that the GID presented in this study is only one interpretation of what a GID could look like and is not necessarily the best representation of the GID's potential.

5.1.3 Future Research

As North (2006) has noted, there are many methods for measuring insight when evaluating a visualization. This study engaged one of those methods in the form of a benchmarked usability test. Future research evaluating GIDs will be able to expand on this study by increasing the sample size to allow for more rigorous statistical analysis and by incorporating additional qualitative assessments, including participants' opinions on the challenges or benefits of a visualization and unstructured interviews where participants use the GID and report their findings without any prompts or guidance.

Aside from adjustments in the methods used to evaluate the GID, future research could also explore other geographic themes within the GID application. The GID used in this study was designed to help assess grassland vegetative development, but the discipline of geography has a multitude of issues that it seeks to address that are well suited for a dashboard visualization. Future research could be conducted to see if performance metrics when using a GID differ among application themes (e.g., social justice, population migration, soil erosion) within the many subfields of geography.

5.1.4 Broader Impacts

This study defines a GID as a spatially-explicit variant within the nascent field of information dashboard research that is currently dominated by applications within the fields of medicine, computer science, and business. The assessment of GID performance here builds on past research in diverse areas including dashboards, GIS, visualization, geographic education, and Cognitive Fit Theory to promote the unique cognitive processes and tasks performed by those in the spatial sciences. This study also represents one of the first evaluations of a dashboard-style visualization outside of the medical field and may help inform other spatial researchers in future investigations.

The results presented here revealed the importance of spatial thinking ability as a key factor influencing participant performance in geographic tasks. The trends observed in task completion time, accuracy, and IES suggest dashboards do have a positive impact on performance across most spatial thinking classes and may be an important interaction term with spatial thinking ability. These trends, along with new usability and utility assessment warrant further investigation and underscores the role and critical nature of geographic education for 21st-century decision-makers.

While the results of this study proved significant, the methodological approach of this study holds great value in and of itself. There is high variability in the methods used to evaluate geographic visualizations, and this variability has led to inconsistent results. This study demonstrates the importance of including time as a dependent variable, instead of solely assessing performance through accuracy. This allowed for a more complete understanding of performance and for the incorporation of the Inverse Efficiency Score. Additionally, this study used three independent variables. Many user studies evaluating visualizations only assess the influence of one independent variable, the type of presentation used. However, this study recognized that when assessing a *geographic* visualization, a user's spatial thinking ability could play a vital role in performance and was thus used as an independent variable. This thesis marks the first study to consider spatial thinking ability as a variable when evaluating a geographic visualization and proves why the inclusion of that measure is so valuable. The experiment also included the independent variable task type. Cognitive Fit Theory has used the classification of tasks to further the understanding of user performance when comparing different representations of data, but this has not been extended to geographic visualization research. The incorporation of this variable allowed for tasks to be categorized by cognitive and spatial thinking difficulty and led to a more holistic understanding of user performance.

By developing a methodological approach that was based on research from multiple fields of study, a deeper understanding of the effectiveness of the GID was achieved. The methods demonstrated here can be easily applied to future studies of dashboards and other geographic visualizations, and the benefits of using this method are evident in the results of this research.

5.2 Conclusion

As visualizations continue to be developed, either within or outside of academia, it is important for those visualizations to be rigorously evaluated for their effectiveness. Research on the evaluation of presentations informs design decisions and ensures that visualizations are reaching their fullest potential. Without this research, it is unknown whether or not a visualization is truly useful and effective in communicating the necessary information to the user. Information dashboards have been developing over the past three decades in medicine and business management, but there is minimal scholarly literature evaluating the dashboard visualization to help guide their design and ensure their effectiveness in communicating information.

This research applied the concepts of information dashboards to geography by developing a GID, then evaluated its ability to promote faster and more accurate completion of spatial thinking tasks. The evaluation of the GID was accomplished through a thorough literature review, pulling research from multiple disciplines, to show why a GID was beneficial to a variety of fields and to provide a framework for evaluating the GID. Results of a controlled experiment, comparing the GID to a non-integrated, tab-based presentation format, showed that the GID allowed participants to answer questions slightly faster and more accurately, but also revealed that spatial thinking ability was the only significant factor in determining the accuracy of responses, regardless of the visualization used.

This research is one of the first studies evaluating the information dashboard visualization for effectiveness and provides a specific framework for future research evaluating the GID. This study provides evidence of the potential dashboards have as a decision support tool and makes the case for including spatial thinking ability in future geovisualization research, regardless of the specific type of visualization being evaluated. The factors controlling the effectiveness of information dashboards in general, and geographic information dashboards in particular, must be better understood in order for dashboard-style visualizations to graduate from interesting novelty to essential decision-making tool.

References

Albert, W.S., and R.G. Golledge. 1999. The use of spatial cognitive abilities in geographic information systems: the map overlay operation. *Transactions in GIS* 3(1): 7-21.

Anselin, L. 2000. Computing environments for spatial data analysis. *Journal of Geographical Systems* 2:201-220.

Anselin, L., Y.W. Kim, and I. Syabri. 2004. Web-based analytical tools for the exploration of spatial data. *Journal of Geographical Systems* 6:197-218. doi:10.1007/s10109-004-0132-5.

Anselin, L., I. Syabri, and Y. Kho. 2006. GeoDa: An introduction to spatial data analysis. *Geographical Analysis* 38:5-22.

Bass, W.M., and R.D. Blanchard. 2011. Examining geographic visualization as a technique for individual risk assessment. *Applied Geography* 31(1): 53-63.

Battersby, S., R.G. Golledge, and M.J. Marsh. 2006. Incidental learning of geospatial concepts across grade level: map overlay. *Journal of Geography* 105(4): 139-146.

Bloom, B.S. 1956. *Taxonomy of Educational Objectives: the Classification of Educational Goals.* New York: D. McKay Co., Inc.

Brewer, C.A. 2005. *Designing Better Maps: A Guide for GIS Users*. Redlands, California: ESRI Press.

Bruyer, R. and M. Brysbaert. 2011. Combining speed and accuracy in cognitive psychology: is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and percentage of errors (PE)?. *Psychologica Belgica* 51(1): 5-13.

Chakravarti, I.M., R.G. Laha, and J. Roy. 1967. *Handbook of Methods of Applied Statistics.* New York: Wiley.

Costa, A. L. 2001. Teacher behaviors that enable student thinking. In *Developing Minds: A Resource Book for Teaching Thinking,* ed. A. L. Costa, 359-369. Alexandria, Virginia: Association for Supervision and Curriculum Development.

Creyer, E.H., J.R. Bettman, and J.W. Payne. 1990. The impact of accuracy and effort feedback and goals on adaptive decision behavior. *Journal of Behavioral Decision Making* 3(1): 1-16.

Cutler, M.E. 1998. The effects of prior knowledge on children's ability to read static and animated maps. MS Thesis. University of South Carolina.

Dennis, A. R., and T. A. Carte. 1998. Using Geographical Information Systems for Decision Making: Extending Cognitive Fit Theory to Map-Based Presentations. *Information Systems Research* 9(2): 194-203.

Dent, B.D., J.S. Torguson, and T.W. Hodler. 2009. *Cartography: Thematic Map Design.* New York: McGraw-Hill.

DiBiase, D., A. M. MacEachren, J. B. Krygier, and C. Reeves. 1992. Animation and the Role of Map Design in Scientific Visualization. *Cartography and Geographic Information Systems* 19(4): 201-214.

Dodge, M., M. McDerby, and M. Turner. 2008. The power of geographic visualizations. In *Geographic Visualization: Concepts, Tools, and Applications,* ed. M. Dodge, M. McDerby, and M. Turner, 1-10. Chichester: John Wiley & Sons Ltd.

Dolan, J.G., P.J. Veazie, and A.J. Russ. 2013. Development and initial evaluation of a treatment decision dashboard. *BMC Medical Informatics and Decision Making* 13(1): 51.

Dunn, S., A.E. Sprague, J.M. Grimshaw, I.D. Graham, M. Taljaard, D. Fell, W.E. Peterson, E. Darling, J. Harrold, G.M. Smith, J. Reszel, A. Lanes, C. Truskoski, J. Wilding, D. Weiss, and M. Walker. 2016. A mixed methods evaluation of the maternal-newborn dashboard in Ontario: dashboard attributes, contextual factors, and facilitators and barriers to use: a study protocol. *Implementation Science* 11(59).

Dykes, J., A. M. MacEachren, and M.-J. Kraak. 2005. Exploring geovisualization. In *Exploring Geovisualization,* ed. J. Dykes, A. M. MacEachren, and M.-J. Kraak, 3-19. Kidlington, United Kingdon: Elsevier Ltd.

Elwood, S. 2010. Geographic information science: visualization, visual methods, and the geoweb. *Progress in Human Geography* 35(3): 401-408.

Few, S. 2006. *Information Dashboard Design: The Effective Visual Communication of Data.* Beijing: O'Reilly.

Flannery, J. J. 1971. The relative effectiveness of some common graduated point symbols in the presentation of quantitative data. *The Canadian Cartographer* 8(2): 96-109.

Gardner, H. 1983. *Frames of Mind: the theory of multiple intelligences*. New York: Basic Books.

Gersmehl, P. 2005. *Teaching Geography*. New York: The Guilford Press.

Gersmehl, P. J., and C. A. Gersmehl. 2007. Spatial thinking by young children: neurologic evidence for early development and "educability". *Journal of Geography* 106(5): 181-191.

Golledge, R. G. 1992. Do people understand spatial concepts: the case of first-order primitives. *Lecture Notes in Computer Science* 639: 1-21.

———. 1995. Primitives of Spatial Knowledge. *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems.*

———. 2002. The Nature of Geographic Knowledge. *Annals of the Association of American Geographers* 91(1): 1-14.

Golledge, R. G., M. Marsh, and S. Battersby. 2008. Matching Geospatial Concepts with Geographic Educational Needs. *Geographical Research* 46(1): 85-98.

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, 69(4), 211–221.

Griffin, A.L., A.M. MacEachren, F. Hardisty, E. Steiner, and B. Li. 2006. A comparison of animated maps with static small-multiple maps for visually identifying space-time clusters. *Annals of the Association of American Geographers* 96(4): 740-753.

Harrower, M. and S. Fabrikant. 2008. The Role of Map Animation for Geographic Visualization. In *Geographic Visualization: Concepts, Tools, and Applications,* ed. M. Dodge, M. McDerby, and M. Turner, 49-65. Chichester: John Wiley & Sons Ltd.

Hermansen, S. 2010. Teaching Cartography in Academia: A Historical Reflection and Discussion of a 2007 Survey of Canadian Universities. *Cartographica* 45(1): 5-18.

Jarvenpaa, S. and G. Dickson. 1988. Graphics and managerial decision making: research-based guidelines. *Communications of the AMC* 31(6): 764-774.

Jo, I. and S. W. Bednarz. 2009. Evaluating Geography Textbook Questions from a Spatial Perspective: Using Concepts of Space, Tools of Representation, and Cognitive Processes to Evaluate Spatiality. *Journal of Geography* 108: 4-13.

Jo, I., S. Bednarz, and S. Metoyer. 2010. Selecting and designing questions to facilitate spatial thinking. *Geography Teacher* 7(2): 49-55.

Jo, I., J.E. Hong, and K. Verma. 2016. Facilitating spatial thinking in world geography using Web-based GIS. *Journal of Geography in Higher Education* 40(3): 442-459.

Karnatak, H.C., R. Shukla, V.K. Sharma, Y.V.S. Murthy, and V. Bhanumurthy. 2012. Spatial mashup technology and real time data integration in geo-web application using open source GIS – a case study for disaster management. *Geocarto International* 27(6): 499-514.

Kavanaugh, M., P. Park, and K. Davis. 2015. Validation of the intensive care unit early warning dashboard: quality improvement utilizing a retrospective case control. *Chest* 148(4): 245A-245A.

Kerski, J.J. 2008. The world at the student's fingertips. In *Digital Geography: Geo-spatial Technologies in the Social Studies Classroom*, ed. A.J. Milson and M. Alibrandi, 119-134. Charlotte, North Carolina: Information Age.

Kim, M. and R. Bednarz. 2013. Development of critical spatial thinking through GIS learning. *Journal of Geography in Higher Education* 3793): 350-366.

Koussoulakou, A. and M.-J. Kraak. 1992. Spatia-temporal maps and cartographic communication. *The Cartographic Journal* 29(2): 101-108.

Koylu, C. and D. Guo. 2017. Design and evaluation of line symbolizations for origin-destination flow maps. *Information Visualization* 16(4): 309-331.

Kraak, M.-J. 2004. The role of the map in a Web-GIS environment. *Journal of Geographical Systems* 6(2): 83-93.

Lee, J. and R. Bednarz. 2009. Effect of GIS learning on spatial thinking. *Journal of Geography in Higher Education* 33(2): 367-390.

Lee, J. and R. Bednarz. 2012. Components of spatial thinking: evidence from a spatial thinking ability test. *Journal of Geography* 111(1): 15-26.

MacEachren, A.M., and M.-J. Kraak. 2001. Research Challenges in Geovisualization. *Cartography and Geographic Information Science* 28(1): 3-12.

McCormick, B.H., T.A. DeFanti, and M.D. Brown. 1987. Visualization in scientific computing – a synopsis. *IEEE Computer Graphics and Applications* 7(4): 61-70.

Mead, R. 1988. *The Design of Experiments.* Cambridge, New York: Cambridge University Press.

Monmonier, M. 1989. Geographic brushing: enhancing exploratory analysis of the scatterplot matrix. *Geographical Analysis* 21(1): 81-84.

Moseley, D., J. Elliott, M. Gregson, and S. Higgins. 2005. Thinking skills frameworks for use in education and training. *British Educational Research Journal* 31(3): 367-390.

Montello, D.R., K.L. Lovelace, R.G. Golledge, and C M. Self. 1999. Sex-related differences and similarities in geographic and environmental spatial abilities. *Annals of the Association of American Geographers* 89(3): 515-534.

Murphy, A. 2014. Geography's Crosscutting Themes: Golden Anniversary Reflections on "The Four Traditions of Geography". *Journal of Geography* 113: 181-188.

National Research Council. 1997. *Rediscovering Geography*. Washington, DC: National Academy Press.

National Research Council. 2006. *Learning to think spatially: GIS as a support system in the K-12 curriculum.* Washington, DC: National Academies Press.

North, C. 2006. Toward measuring visualization insight. *IEEE Computer Graphics and Applications* 26(3): 6-9.

Pattinson, W.D. 1964. The Four Traditions of Geography. *Journal of Geography* 63(5): 211-216.

Rinner, C. 2007. A geographic visualization approach to multi-criteria evaluation of urban quality of life. *International Journal of Geographical Information Science* 21(8): 907-919.

Robinson, A. 2017. Geovisual analytics. *The Geographic Information Science & Technology Body of Knowledge*, ed. J. P. Wilson. University Consortium for Geographic Information Science.

Rodgers, P. 2005. Graph Drawing Techniques for Geographic Visualization. In *Exploring Geovisualization*, ed. J. Dykes, A.M. MacEachren, and M.-J. Kraak, 143-158. Kidlington, United Kingdom: Elsevier Ltd.

Roth, R. 2012. Cartographic interaction primitives: framework and synthesis. *The Cartographic Journal* 49(4): 376-395.

———. 2013a. Interactive maps: what we know and what we need to know. *Journal of Spatial Information Science* 2013(6): 59-115.

———. 2013b. An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. *IEEE Transactions on Visualization and Computer Graphics* 19(12): 2356-2365.

———. 2015. Interactivity and cartography: a contemporary perspective on user interface and user experience design from geospatial professionals. *Cartographica: The International Journal for Geographic Information and Geovisualization* 50(2): 94-115.

Roth, R., K. Ross, and A. MacEachren. 2015. User-centered design for interactive maps: a case study in crime analysis. *ISPRS International Journal of Geo-Information* 4(1): 262-301.

Speier, C. 2006. The influence of information presentation formats on complex task decision-making performance. *International Journal of Human-Computer Studies* 64: 1115-1131.

Stahl, R.J. and G.T. Murphy. 1981. The domain of cognition: an alternative to Bloom's cognitive domain within the framework of an information processing model.

Tobón, C. 2005. Evaluating Geographic Visualization Tools and Methods: An Approach and Experiment Based upon User Tasks. In *Exploring Geovisualization*, ed. J. Dykes, A. M. MacEachren, and M.-J. Kraak, 645-666. Kidlington, United Kingdom: Elsevier Ltd.

Vessey, I. 1991. Cognitive Fit: A Theory-Based Analysis of the Graphs Versus Tables Literature. *Decision Sciences* 22(2): 219-240.

Wood, J., S. Kirschenbauer, J. Dollner, A. Lopes, and L. Bodum. 2005. Using 3D in Visualization. In *Exploring Geovisualization*, ed. J. Dykes, A. M. MacEachren, and M.-J. Kraak, 295-312. Kidlington, United Kingdom: Elsevier Ltd.

Yigitbasioglu, O.M., and O. Velcu. 2012. A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems* 13: 41-59.

Appendix A - RShiny Script for Dashboard Visualization

```
library(shiny)

source("global.R")

ui <- fluidPage(
  theme = shinytheme("cerulean"),
  titlePanel("Fort Riley Range and Training Land Assessment"),

  fluidRow(

    column(4,
      selectInput(inputId = "date", "View Image from Date:",
        pics <- c(
         "January 1, 2012"="mi120101",
          "January 17, 2012"="mi120117",
          "February 2, 2012"="mi120202",
          "February 18, 2012"="mi120218",
          "March 6, 2012"="mi120306",
          "March 22, 2012"="mi120322",
          "April 7, 2012"="mi120407",
          "April 23, 2012"="mi120423",
          "May 9, 2012"="mi120509",
          "May 25, 2012"="mi120525",
          "June 10, 2012"="mi120610",
          "June 26, 2012"="mi120626",
          "July 12, 2012"="mi120712",
          "July 28, 2012"="mi120728",
          "August 13, 2012"="mi120813",
          "August 29, 2012"="mi120829",
          "September 14, 2012"="mi120914",
          "September 30, 2012"="mi120930",
          "October 16, 2012"="mi121016",
          "November 1, 2012"="mi121101",
          "November 17, 2012"="mi121117",
          "December 3, 2012"="mi121203",
          "December 19, 2012"="mi121219"),
          multiple = FALSE)
    ),

    column(4,
      selectInput(
        inputId = "date2",
        label = "Calculate NDVI Change To:",
        choices = pics,
        multiple = FALSE)
    ),

    column(4,
          selectInput(inputId = "index", "View Phenology from January 1 to:",
                      pics <- c(
                        "January 1, 2012"=1,
                        "January 17, 2012"=2,
                        "February 2, 2012"=3,
                        "February 18, 2012"=4,
                        "March 6, 2012"=5,
                        "March 22, 2012"=6,
                        "April 7, 2012"=7,
                        "April 23, 2012"=8,
                        "May 9, 2012"=9,
                        "May 25, 2012"=10,
```

```
                              "May 25, 2012"=10,
                              "June 10, 2012"=11,
                              "June 26, 2012"=12,
                              "July 12, 2012"=13,
                              "July 28, 2012"=14,
                              "August 13, 2012"=15,
                              "August 29, 2012"=16,
                              "September 14, 2012"=17,
                              "September 30, 2012"=18,
                              "October 16, 2012"=19,
                              "November 1, 2012"=20,
                              "November 17, 2012"=21,
                              "December 3, 2012"=22,
                              "December 19, 2012"=23),
                            multiple = FALSE)
    )
  ),

  fluidRow(
    column(4,
      leafletOutput("map"),
      htmlOutput("mapavg")
    ),

    column(4,
      leafletOutput("diff"),
      htmlOutput("diffavg")
    ),

    column(4,
      plotOutput("pheno")
    )
  ),

  fluidRow(
    column(4,
      plotOutput("hist")

    ),

    column(4,
      plotOutput("diffhist")
    ),

    column(4,
      plotOutput("weather")
    )
  )
)

server <- function(input, output) {

  weatherData <- reactive({
    test <- weatherYear[weatherYear$MODIS %in% seq(from=1, to=max(input$index),by=1),]

  })

```

```r

output$map <- renderLeaflet({
  leaflet() %>% addProviderTiles("CartoDB.Positron") %>%
    setView(-96.73, 39.17, zoom=10) %>%
    addRasterImage(ndviStack[[input$date]], opacity=0.7, col=cols1) %>%
    addPolygons(data=frk, color="black", weight=1, fill=FALSE) %>%
    addLabelOnlyMarkers(frk$long, frk$lat, label=frk$TRNG_ID,
               labelOptions = labelOptions(noHide=T, textOnly=T))%>%
    addPolygons(data=knz, color="black", weight=1, fill=FALSE) %>%
    addDrawToolbar(
      targetGroup="draw",
      position="bottomleft",
      polylineOptions=FALSE,
      polygonOptions=FALSE,
      circleOptions=FALSE,
      markerOptions=FALSE,
      singleFeature=TRUE) %>%
    addLayersControl(
      overlayGroups=c("draw"),
      options = layersControlOptions(collapsed=FALSE),
      position="bottomright"
    )
})

output$mapavg <- renderText({
  rasterobj = mapRaster(input)
  avgpx <- cellStats(rasterobj, stat='mean', na.rm=TRUE)
  paste("<B>Avg Pixel Value: </B>", format(round(avgpx, 2), nsmall=2))
})

output$hist <- renderPlot({
  rasterobj = mapRaster(input)
  hist(rasterobj, main="", xlab="NDVI (8-bit DN)", breaks=brks1, col=cols1)
})

output$pheno <- renderPlot({
  graph.pheno <- pheno.data[1:input$index,]
  ggplot() +
    geom_area(aes(y = cellnorms, x = index), data=pheno.data,
           fill = "lightgreen", alpha  = 0.5) +
    geom_point(aes(y = cellstats, x = index), data=graph.pheno,
               color = "darkgreen", pch = 15, cex = 2.5) +
    geom_line(aes(y = cellstats, x = index), data=graph.pheno, color = "darkgreen", lwd = 1) +
    labs(y = "NDVI", x = "Month") +
    coord_cartesian(ylim = c(80, 190), xlim = c(0.5, 23.5)) +
    scale_x_continuous(expand = c(0, 0),
                       breaks = seq(1.9,23,1.9),
                       labels = c("J", "F", "M", "A", "M", "J", "J", "A", "S", "O", "N", "D")) +
    theme_bw()
})

output$diff <- renderLeaflet({
  leaflet() %>% addProviderTiles("CartoDB.Positron") %>%
    setView(-96.73, 39.17, zoom=10) %>%
    addRasterImage(overlay(x=ndviStack[[input$date]], y=ndviStack[[input$date2]],
                           fun=percDiff, opacity=0.7, col=cols2)) %>%
    addPolygons(data=frk, color="black", weight=1, fill=FALSE) %>%
```

```
      addPolygons(data=frk, color="black", weight=1, fill=FALSE) %>%
      addLabelOnlyMarkers(frk$long, frk$lat, label=frk$TRNG_ID,
                          labelOptions = labelOptions(noHide=T, textOnly=T))%>%
      addPolygons(data=knz, color="black", weight=1, fill=FALSE) %>%
      addDrawToolbar(
        targetGroup="draw",
        position="bottomleft",
        polylineOptions=FALSE,
        polygonOptions=FALSE,
        circleOptions=FALSE,
        markerOptions=FALSE,
        singleFeature=TRUE) %>%
      addLayersControl(
        overlayGroups=c("draw"),
        options = layersControlOptions(collapsed=FALSE),
        position="bottomright"
      )
  })

  output$diffavg <- renderText({
    rasterobj = diffRaster(input)
    avgpx <- cellStats(rasterobj, stat='mean', na.rm=TRUE)
    paste("<B>Avg Percent Change: </B>", format(round(avgpx, 2), nsmall=2))
  })

  output$diffhist <- renderPlot({
    rasterobj = diffRaster(input)
    hist(rasterobj, col=cols2, breaks=brks2, xlab="Percent Change in NDVI", main="", freq=TRUE)
  })

  output$weather <- renderPlot({
    ggplot(weatherData(), aes(x=DOY)) +
      geom_bar(aes(y = CompPrecip*15), stat = "identity", color = "blue") +
      geom_ribbon(aes(ymin = TempMin, ymax = TempMax), alpha = 0.5, fill = "indianred2") +
      geom_line(aes(y = TempAvg), color = "black") +
      #geom_bar(aes(y = Precip*50), stat = "identity", color = "blue") +
      labs(y = "Mean Daily Temperature (F)", x = "Month") +
      coord_cartesian(ylim = c(0,105), xlim = c(1, 365)) +
      #scale_y_continuous(sec.axis = sec_axis(~./50, name="Precipitation (in)")) +
      scale_y_continuous(sec.axis = sec_axis(~./15, name="Precipitation (in)")) +
      scale_x_continuous(expand = c(0, 0),
                         breaks = c(15,45,75,105,135,165,195,228,258,288,320,350),
                         labels = c("J", "F", "M", "A", "M", "J", "J", "A", "S", "O", "N", "D")) +
      theme_bw()
  })

}

shinyApp(ui = ui, server = server)
```

```
#load Save current document (Ctrl+S) ries
librar
library(leaflet)
library(raster)
library(rgdal)
library(raster)
library(rasterVis)
library(dplyr)
library(tidyr)
library(zoo)
library(DataCombine)
library(ggplot2)
library(gridExtra)
library(magrittr)
library(sp)
library(leaflet.extras)
library(shinythemes)

#set working directory
#setwd("Z:\\Projects\\Fischer_Amariah\\Shiny\\SampleGrids")

#create list of grids and create raster stack
list <- list.files('SampleGrids', pattern = glob2rx('*.tif'), full.names = TRUE)
ndviStack <- stack(list)
cellstats <- cellStats(ndviStack, "mean")

#create NDVI Normal stack
#setwd("Z:\\Projects\\Fischer_Amariah\\Shiny\\NormalGrids")
list2 <- list.files('NormalGrids', pattern = glob2rx('*.tif'), full.names = TRUE)
normStack <- stack(list2)
cellnorms <- cellStats(normStack, "mean")

#create data frame combining annual and normal phenometrics for entire year
pheno.data <- data.frame(cellstats=cellstats, cellnorms=cellnorms)
pheno.data$index <- seq(1:23)

#function to calculate NDVI difference image
percDiff <- function(x, y) {
  pdif <- ((y - x) / 250) * 100
    return(pdif)
}


#load shapefiles
#frk <- spTransform(readOGR("Z:\\Projects\\Fischer_Amariah\\Shiny\\frkviewer\\Data\\Shapefiles",
  "riley_bound"), CRS("+proj=longlat +datum=WGS84"))
#knz <- spTransform(readOGR("Z:\\Projects\\Fischer_Amariah\\Shiny\\frkviewer\\Data\\Shapefiles",
  "konza_bound"), CRS("+proj=longlat +datum=WGS84"))
frk <- readOGR("Shapefiles", "riley_bound")
knz <- readOGR("Shapefiles", "konza_bound")

#controls for graph aesthetics
brks1 <- seq(0, 250, by=10)
cols1 <- colorRampPalette(c("saddlebrown", "white", "green4"))(length(brks1)-1)
#brks2 <- c(-100,-20,-15,-10,-5,5,10,15,20,100)
#cols2 <- colorRampPalette(brewer.pal(11,"RdYlGn"))(length(brks2 length(x)
brks2 <- seq(-100,100, by=10)
cols2 <- c("#A50026","#A50026","#A50026","#A50026","#A50026","#A50026","#A50026","#DE3F2E",
           "#F88D51","#FDD380","#FFFFBF","#CCE982","#86CB66","#2DA154","#006837","#006837",
           "#006837","#006837","#006837","#006837","#006837")
```

```
            "#006837","#006837","#006837","#006837","#006837")

#import weather datafile and create variable of just observation dates
#setwd("Z:\\Projects\\Fischer_Amariah\\Shiny\\Weather")
w2 <- read.csv(file="Weather/NewManhattanWeather2001_2017.csv",header=TRUE,
               sep=",", na.strings=c(NA, "NA", ""))

#subset datafile by variable and rename variables
myvars <- c("Month", "Day", "Year", "DOY", "AirTempMax.F.", "AirTempMin.F.", "Precip.In.",
            "GDDday", "GDDaccum", "MODIS_No", "CompPrecip")
w <- w2[myvars]
names(w) <- c("Month", "Day", "Year", "DOY", "TempMax", "TempMin", "Precip", "DailyGDD",
              "CumGDD", "MODIS", "CompPrecip")
w["TempAvg"] <- (w$TempMax + w$TempMin) /2
weatherYear <- w[ which(w$Year == 2012),]

#custom functions for drawing the polygon
convertCoords = function(x, y, crsstring){
  umt = SpatialPoints(cbind(x,y), proj4string=crs(crsstring))
  ll = spTransform(umt, CRS('+proj=longlat'))
  return(attributes(extent(ll)))
}

convertRasterBounds = function(rasterobj){
  attrs = attributes(extent(rasterobj))
  crsstring = projection(rasterobj)
  min = convertCoords(attrs$xmin, attrs$ymin, crsstring)
  max = convertCoords(attrs$xmax, attrs$ymax, crsstring)
  return(list(x1 = min$xmin, x2 = max$xmax, y1 = max$ymax, y2 = min$ymin))
}

cropRasterToRectPoly = function(rasterobj, polycoords){
  rastLL = convertRasterBounds(rasterobj)
  polyLL = list(x1 = polycoords[[1]][[1]][1], x2 = polycoords[[3]][[1]][1],
                y1 = polycoords[[3]][[2]][1], y2 = polycoords[[1]][[2]][1])
  nrows = nrow(rasterobj)
  ncols = ncol(rasterobj)

  degColWidth = abs(rastLL$x2 - rastLL$x1) / ncols
  degRowHeight = abs(rastLL$y2 - rastLL$y1) / nrows
  if( rastLL$x1 > rastLL$x2 || polyLL$x1 > polyLL$x2 || rastLL$y1 < rastLL$y2 || polyLL$y1 < polyLL$y2 ){
    stop("The coords aren't in the order we expected so the maths will be wrong")
  }
  if( polyLL$x1 > rastLL$x2 || polyLL$x2 < rastLL$x1 || polyLL$y1 < rastLL$y2 || polyLL$y2 > rastLL$y1 ){
    stop('Poly must overlap raster')
  }
  #min row number (ratio of polyLL x1 to rastLL x1<>x2 = ratio of x1 to 1<>nrows)
  x1 = ifelse( polyLL$x1 < rastLL$x1, 1, ceiling((polyLL$x1 - rastLL$x1) / degColWidth) )
  x2 = ifelse( polyLL$x2 > rastLL$x2, ncols, floor((polyLL$x2 - rastLL$x1) / degColWidth) )
  y1 = ifelse( polyLL$y1 > rastLL$y1, 1, ceiling((rastLL$y1 - polyLL$y1) / degRowHeight) )
  y2 = ifelse( polyLL$y2 < rastLL$y2, nrows, floor((rastLL$y1 - polyLL$y2) / degRowHeight) )
  return(crop(rasterobj, extent(rasterobj, y1, y2, x1, x2)))
}

#code that makes other stuff happen more easily
diffRaster <- function(input){
  diff <- overlay(x=ndviStack[[input$date]], y=ndviStack[[input$date2]], fun=percDiff)
```


```
  diff <- overlay(x=ndviStack[[input$date]], y=ndviStack[[input$date2]], fun=percDiff)
```


```
  rasterobj = diff
  if( isTruthy(input$diff_draw_stop) ){
    polygon_coordinates = input$diff_draw_new_feature$geometry$coordinates[[1]]
    if( is.null(polygon_coordinates) ){
      stop('Must draw a poly')
    }
    rasterobj = cropRasterToRectPoly(rasterobj, polygon_coordinates)
  }
  return(rasterobj)
}

mapRaster <- function(input){
  rasterobj = ndviStack[[input$date]]
  if( isTruthy(input$map_draw_stop) ){
    polygon_coordinates = input$map_draw_new_feature$geometry$coordinates[[1]]
    if( is.null(polygon_coordinates) ){
      stop('Must draw a poly')
    }
    rasterobj = cropRasterToRectPoly(rasterobj, polygon_coordinates)
  }
  return(rasterobj)
}
```

Appendix B - RShiny Script for Tabbed Visualization

```
library(shiny)

source("globalTabbed.R")

ui <- fluidPage(
  theme = shinytheme("cerulean"),
  headerPanel("Fort Riley Range and Training Land Assessment"),

  mainPanel(
    tabsetPanel(
      tabPanel("Map",

      selectInput(inputId = "date", "View Image from Date:",
        pics <- c(
          "January 1, 2012"="mi120101",
          "January 17, 2012"="mi120117",
          "February 2, 2012"="mi120202",
          "February 18, 2012"="mi120218",
          "March 6, 2012"="mi120306",
          "March 22, 2012"="mi120322",
          "April 7, 2012"="mi120407",
          "April 23, 2012"="mi120423",
          "May 9, 2012"="mi120509",
          "May 25, 2012"="mi120525",
          "June 10, 2012"="mi120610",
          "June 26, 2012"="mi120626",
          "July 12, 2012"="mi120712",
          "July 28, 2012"="mi120728",
          "August 13, 2012"="mi120813",
          "August 29, 2012"="mi120829",
          "September 14, 2012"="mi120914",
          "September 30, 2012"="mi120930",
          "October 16, 2012"="mi121016",
          "November 1, 2012"="mi121101",
          "November 17, 2012"="mi121117",
          "December 3, 2012"="mi121203",
          "December 19, 2012"="mi121219"),
          multiple = FALSE),
      leafletOutput("map"),
      htmlOutput("mapavg")
      ),

      tabPanel("Difference Map",
        fluidRow(
      column(6,
        selectInput(
        inputId = "date2",
        label = "Calculate NDVI Change From:",
        choices = pics,
        multiple = FALSE)),

      column(6,
        selectInput(
        inputId = "date3",
        label = "To:",
        choices = pics,
        multiple = FALSE))
      ),

      leafletOutput("diff"),
```

```
    leafletOutput("diff"),
    htmlOutput("diffavg")
    ),

tabPanel("Map Histogram",
         selectInput(
           inputId = "date6",
           label = "Calculate NDVI Change From:",
           choices = pics,
           multiple = FALSE),

    plotOutput("hist")
),


  tabPanel("Difference Map Histogram",
           fluidRow(
             column(6,
                    selectInput(
                      inputId = "date4",
                      label = "Calculate NDVI Change From:",
                      choices = pics,
                      multiple = FALSE)),

             column(6,
                    selectInput(
                      inputId = "date5",
                      label = "To:",
                      choices = pics,
                      multiple = FALSE))
           ),

    plotOutput("diffhist")
  ),

tabPanel("Phenology",

         selectInput(inputId = "index", "View Phenology from January 1 to:",
                     pics <- c(
                       "January 1, 2012"=1,
                       "January 17, 2012"=2,
                       "February 2, 2012"=3,
                       "February 18, 2012"=4,
                       "March 6, 2012"=5,
                       "March 22, 2012"=6,
                       "April 7, 2012"=7,
                       "April 23, 2012"=8,
                       "May 9, 2012"=9,
                       "May 25, 2012"=10,
                       "June 10, 2012"=11,
                       "June 26, 2012"=12,
                       "July 12, 2012"=13,
                       "July 28, 2012"=14,
                       "August 13, 2012"=15,
                       "August 29, 2012"=16,
                       "September 14, 2012"=17,
                       "September 30, 2012"=18,
```

```
                         "September 30, 2012"=18,
                         "October 16, 2012"=19,
                         "November 1, 2012"=20,
                         "November 17, 2012"=21,
                         "December 3, 2012"=22,
                         "December 19, 2012"=23),
                       multiple = FALSE),
           plotOutput("pheno")
  ),

    tabPanel("Weather",

             selectInput(inputId = "index2", "View Weather from January 1 to:",
                         pics <- c(
                           "January 1, 2012"=1,
                           "January 17, 2012"=2,
                           "February 2, 2012"=3,
                           "February 18, 2012"=4,
                           "March 6, 2012"=5,
                           "March 22, 2012"=6,
                           "April 7, 2012"=7,
                           "April 23, 2012"=8,
                           "May 9, 2012"=9,
                           "May 25, 2012"=10,
                           "June 10, 2012"=11,
                           "June 26, 2012"=12,
                           "July 12, 2012"=13,
                           "July 28, 2012"=14,
                           "August 13, 2012"=15,
                           "August 29, 2012"=16,
                           "September 14, 2012"=17,
                           "September 30, 2012"=18,
                           "October 16, 2012"=19,
                           "November 1, 2012"=20,
                           "November 17, 2012"=21,
                           "December 3, 2012"=22,
                           "December 19, 2012"=23),
                         multiple = FALSE),
      plotOutput("weather")
    )


)))

server <- function(input, output) {

  weatherData <- reactive({
    test <- weatherYear[weatherYear$MODIS %in% seq(from=1, to=max(input$index2),by=1),]

  })

  output$map <- renderLeaflet({
    leaflet() %>% addProviderTiles("CartoDB.Positron") %>%
      setView(-96.73, 39.17, zoom=10) %>%
      addRasterImage(ndviStack[[input$date]], opacity=0.7, col=cols1) %>%
      addPolygons(data=frk, color="black", weight=1, fill=FALSE) %>%
      addLabelOnlyMarkers(frk$long, frk$lat, label=frk$TRNG_ID,
```

```
      addLabelOnlyMarkers(frk$long, frk$lat, label=frk$TRNG_ID,
                          labelOptions = labelOptions(noHide=T, textOnly=T))%>%
      addPolygons(data=knz, color="black", weight=1, fill=FALSE) %>%
      addDrawToolbar(
        targetGroup="draw",
        position="bottomleft",
        polylineOptions=FALSE,
        polygonOptions=FALSE,
        circleOptions=FALSE,
        markerOptions=FALSE,
        singleFeature=TRUE) %>%
      addLayersControl(
        overlayGroups=c("draw"),
        options = layersControlOptions(collapsed=FALSE),
        position="bottomright"
      )
  })

  output$mapavg <- renderText({
    rasterobj = mapRaster(input)
    avgpx <- cellStats(rasterobj, stat='mean', na.rm=TRUE)
    paste("<B>Avg Pixel Value: </B>", format(round(avgpx, 2), nsmall=2))
  })

  output$hist <- renderPlot({
    hist(ndviStack[[input$date6]], main="", xlab="NDVI (8-bit DN)", breaks=brks1, col=cols1)
  })

  output$pheno <- renderPlot({
    graph.pheno <- pheno.data[1:input$index,]
    ggplot() +
      geom_area(aes(y = cellnorms, x = index), data=pheno.data, fill = "lightgreen", alpha  = 0.5) +
      geom_point(aes(y = cellstats, x = index),
                 data=graph.pheno, color = "darkgreen", pch = 15, cex = 2.5) +
      geom_line(aes(y = cellstats, x = index), data=graph.pheno, color = "darkgreen", lwd = 1) +
      labs(y = "NDVI", x = "Month") +
      coord_cartesian(ylim = c(80, 190), xlim = c(0.5, 23.5)) +
      scale_x_continuous(expand = c(0, 0),
                         breaks = seq(1.9,23,1.9),
                         labels = c("J", "F", "M", "A", "M", "J", "J", "A", "S", "O", "N", "D")) +
      theme_bw()
  })

  output$diff <- renderLeaflet({
    leaflet() %>% addProviderTiles("CartoDB.Positron") %>%
      setView(-96.73, 39.17, zoom=10) %>%
      addRasterImage(overlay(x=ndviStack[[input$date2]], y=ndviStack[[input$date3]],
                             fun=percDiff, opacity=0.7, col=cols2)) %>%
      addPolygons(data=frk, color="black", weight=1, fill=FALSE) %>%
      addLabelOnlyMarkers(frk$long, frk$lat, label=frk$TRNG_ID,
                          labelOptions = labelOptions(noHide=T, textOnly=T))%>%
      addPolygons(data=knz, color="black", weight=1, fill=FALSE) %>%
      addDrawToolbar(
        targetGroup="draw",
        position="bottomleft",
        polylineOptions=FALSE,
        polygonOptions=FALSE,
```

```
        polygonOptions=FALSE,
        circleOptions=FALSE,
        markerOptions=FALSE,
        singleFeature=TRUE) %>%
      addLayersControl(
        overlayGroups=c("draw"),
        options = layersControlOptions(collapsed=FALSE),
        position="bottomright"
      )
  })

  output$diffavg <- renderText({
    rasterobj = diffRaster(input)
    avgpx <- cellStats(rasterobj, stat='mean', na.rm=TRUE)
    paste("<B>Avg Percent Change: </B>", format(round(avgpx, 2), nsmall=2))
  })

  output$diffhist <- renderPlot({
    rasterobj = overlay(x=ndviStack[[input$date4]], y=ndviStack[[input$date5]], fun=percDiff)
    hist(rasterobj, col=cols2, breaks=brks2, xlab="Percent Change in NDVI", main="")
  })

  output$weather <- renderPlot({
    ggplot(weatherData(), aes(x=DOY)) +
      geom_bar(aes(y = CompPrecip*15), stat = "identity", color = "blue") +
      geom_ribbon(aes(ymin = TempMin, ymax = TempMax), alpha = 0.5, fill = "indianred2") +
      geom_line(aes(y = TempAvg), color = "black") +
      #geom_bar(aes(y = Precip*50), stat = "identity", color = "blue") +
      labs(y = "Mean Daily Temperature (F)", x = "Month") +
      coord_cartesian(ylim = c(0,105), xlim = c(1, 365)) +
      #scale_y_continuous(sec.axis = sec_axis(~./50, name="Precipitation (in)")) +
      scale_y_continuous(sec.axis = sec_axis(~./15, name="Precipitation (in)")) +
      scale_x_continuous(expand = c(0, 0),
                         breaks = c(15,45,75,105,135,165,195,228,258,288,320,350),
                         labels = c("J", "F", "M", "A", "M", "J", "J", "A", "S", "O", "N", "D")) +
      theme_bw()
  })

}

shinyApp(ui = ui, server = server)
```

```r
#load required libraries
library(shiny)
library(leaflet)
library(raster)
library(rgdal)
library(raster)
library(rasterVis)
library(dplyr)
library(tidyr)
library(zoo)
library(DataCombine)
library(ggplot2)
library(gridExtra)
library(magrittr)
library(sp)
library(leaflet.extras)
library(shinythemes)

#set working directory
#setwd("Z:\\Projects\\Fischer_Amariah\\Shiny\\SampleGrids")

#create list of grids and create raster stack
list <- list.files('SampleGrids', pattern = glob2rx('*.tif'), full.names = TRUE)
ndviStack <- stack(list)
cellstats <- cellStats(ndviStack, "mean")

#create NDVI Normal stack
#setwd("Z:\\Projects\\Fischer_Amariah\\Shiny\\NormalGrids")
list2 <- list.files('NormalGrids', pattern = glob2rx('*.tif'), full.names = TRUE)
normStack <- stack(list2)
cellnorms <- cellStats(normStack, "mean")

#create data frame combining annual and normal phenometrics for entire year
pheno.data <- data.frame(cellstats=cellstats, cellnorms=cellnorms)
pheno.data$index <- seq(1:23)

#function to calculate NDVI difference image
percDiff <- function(x, y) {
  pdif <- ((y - x) / 250) * 100
    return(pdif)
}


#load shapefiles
#frk <- spTransform(readOGR("Z:\\Projects\\Fischer_Amariah\\Shiny\\frkviewer\\Data\\Shapefiles",
  "riley_bound"), CRS("+proj=longlat +datum=WGS84"))
#knz <- spTransform(readOGR("Z:\\Projects\\Fischer_Amariah\\Shiny\\frkviewer\\Data\\Shapefiles",
  "konza_bound"), CRS("+proj=longlat +datum=WGS84"))
frk <- readOGR("Shapefiles", "riley_bound")
knz <- readOGR("Shapefiles", "konza_bound")

#controls for graph aesthetics
brks1 <- seq(0, 250, by=10)
cols1 <- colorRampPalette(c("saddlebrown", "white", "green4"))(length(brks1)-1)
#brks2 <- c(-100,-20,-15,-10,-5,5,10,15,20,100)
#cols2 <- colorRampPalette(brewer.pal(11,"RdYlGn"))(length(brks2)-1)
brks2 <- seq(-100,100, by=10)
cols2 <- c("#A50026","#A50026","#A50026","#A50026","#A50026","#A50026","#A50026","#DE3F2E",
           "#F88D51","#FDD380","#FFFFBF","#CCE982","#86CB66","#2DA154","#006837","#006837",
           "#006837","#006837","#006837","#006837","#006837")
```

```
            "#006837","#006837","#006837","#006837","#006837")

#import weather datafile and create variable of just observation dates
#setwd("Z:\\Projects\\Fischer_Amariah\\Shiny\\Weather")
w2 <- read.csv(file="Weather/NewManhattanWeather2001_2017.csv",header=TRUE,
               sep=",", na.strings=c(NA, "NA", ""))

#subset datafile by variable and rename variables
myvars <- c("Month", "Day", "Year", "DOY", "AirTempMax.F.",
            "AirTempMin.F.", "Precip.In.", "GDDday", "GDDaccum", "MODIS_No")
w <- w2[myvars]
names(w) <- c("Month", "Day", "Year", "DOY", "TempMax", "TempMin",
             "Precip", "DailyGDD", "CumGDD", "MODIS")
w["TempAvg"] <- (w$TempMax + w$TempMin) /2
w$MODIS <- as.factor(w$MODIS)
weatherYear <- w[ which(w$Year == 2012),]

#custom functions for drawing the polygon
convertCoords = function(x, y, crsstring){
  umt = SpatialPoints(cbind(x,y), proj4string=crs(crsstring))
  ll = spTransform(umt, CRS('+proj=longlat'))
  return(attributes(extent(ll)))
}

convertRasterBounds = function(rasterobj){
  attrs = attributes(extent(rasterobj))
  crsstring = projection(rasterobj)
  min = convertCoords(attrs$xmin, attrs$ymin, crsstring)
  max = convertCoords(attrs$xmax, attrs$ymax, crsstring)
  return(list(x1 = min$xmin, x2 = max$xmax, y1 = min$ymin, y2 = max$ymax))
}

cropRasterToRectPoly = function(rasterobj, polycoords){
  rastLL = convertRasterBounds(rasterobj)
  polyLL = list(x1 = polycoords[[1]][[1]][1], x2 = polycoords[[3]][[1]][1],
                y1 = polycoords[[1]][[2]][1], y2 = polycoords[[3]][[2]][1])
  nrows = nrow(rasterobj)
  ncols = ncol(rasterobj)

  degColWidth = (rastLL$x2 - rastLL$x1) / ncols
  degRowHeight = (rastLL$y2 - rastLL$y1) / nrows
  if( rastLL$x1 > rastLL$x2 || polyLL$x1 > polyLL$x2 || rastLL$y1 > rastLL$y2 || polyLL$y1 > polyLL$y2 ){
    stop("The coords aren't in the order we expected so the maths will be wrong")
  }
  if( polyLL$x1 > rastLL$x2 || polyLL$x2 < rastLL$x1 || polyLL$y1 > rastLL$y2 || polyLL$y2 < rastLL$y1 ){
    stop('Poly must overlap raster')
  }
  #min row number (ratio of polyLL x1 to rastLL x1<>x2 = ratio of x1 to 1<>nrows)
  x1 = ifelse( polyLL$x1 < rastLL$x1, 1, ceiling((polyLL$x1 - rastLL$x1) / degColWidth) )
  x2 = ifelse( polyLL$x2 > rastLL$x2, ncols, floor((polyLL$x2 - rastLL$x1) / degColWidth) )
  y1 = ifelse( polyLL$y1 < rastLL$y1, 1, ceiling((polyLL$y1 - rastLL$y1) / degRowHeight) )
  y2 = ifelse( polyLL$y2 > rastLL$y2, nrows, floor((polyLL$y2 - rastLL$y1) / degRowHeight) )
  return(crop(rasterobj, extent(rasterobj, y1, y2, x1, x2)))
}

#code that makes other stuff happen more easily
diffRaster <- function(input){
```

```
diffRaster <- function(input){
  diff <- overlay(x=ndviStack[[input$date2]], y=ndviStack[[input$date3]], fun=percDiff)
  rasterobj = diff
  if( isTruthy(input$diff_draw_stop) ){
    polygon_coordinates = input$diff_draw_new_feature$geometry$coordinates[[1]]
    if( is.null(polygon_coordinates) ){
      stop('Must draw a poly')
    }
    rasterobj = cropRasterToRectPoly(rasterobj, polygon_coordinates)
  }
  return(rasterobj)
}

mapRaster <- function(input){
  rasterobj = ndviStack[[input$date]]
  if( isTruthy(input$map_draw_stop) ){
    polygon_coordinates = input$map_draw_new_feature$geometry$coordinates[[1]]
    if( is.null(polygon_coordinates) ){
      stop('Must draw a poly')
    }
    rasterobj = cropRasterToRectPoly(rasterobj, polygon_coordinates)
  }
  return(rasterobj)
}
```

Appendix C - Online Qualtrics Survey Questions

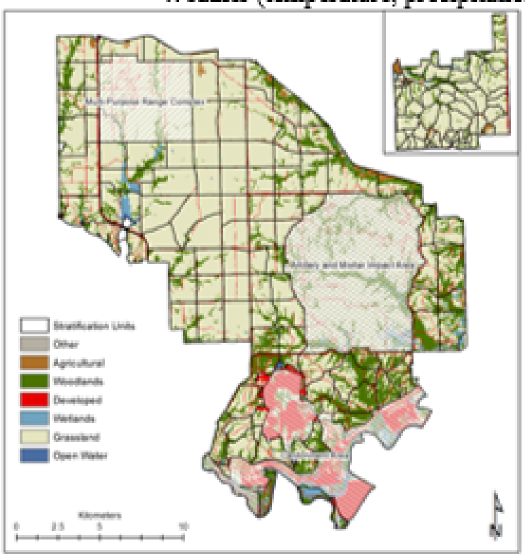
Question Number	Question	Points Possible
1	What is the average pixel value of the image on March 22, 2012?	1
2	Is the average pixel value for March 22, 2012 normal for this time of year? If not, is it higher or lower than normal?	2
3	Hypothesize what could be causing the average pixel value for March 22, 2012 to be normal or abnormal.	2
4	Looking at the data for April 7, 2012, can you identify any training area(s) that might be of concern to a land manager at Fort Riley? If yes, what are the training area numbers? (If you identify multiple areas of concern, choose the area that appears the most concerning)	1
5	How does the average pixel value for April 7, 2012 compare to the average pixel value for the region you identified in the previous question?	1
6	Calculate the NDVI change from April 7 to April 23. What is the approximate difference between the average percent change for the whole image and the region you identified?	1
7	Is the difference in average percent change significant?	1
8	Based on the data for April 7 and April 23, should this region be of concern to a land manager at Fort Riley? Explain.	2
9	Looking at phenology for the whole year, do you find the year 2012 to be close to the normal phenology curve or not? Explain.	2
10	Hypothesize why the year 2012 is or is not similar to the normal phenology curve.	2
11	Between August 13 and August 29 there is a sudden increase in NDVI pixel values. Hypothesize what could have caused this sudden increase.	2
12	Compare and contrast the histogram from December 3 to December 19.	2
13	Do the differences/similarities make sense for this time of year?	2
14	Calculate the NDVI change from December 3 to December 19. Describe the spatial distribution of NDVI percent change for this image.	2
15	Compare and contrast the percent change in Fort Riley to that of Konza from December 3 to December 19.	2
16	Hypothesize why there are similarities/differences between Fort Riley and Konza for this time period.	2
17	Describe the NDVI data from May 2012 to August 2012.	1
18	Compare and contrast the NDVI data from May to August in 2012 to the normal NDVI data from May to August.	1

19	Hypothesize why there are differences/similarities between the normal and 2012 NDVI data from May to August.	2
20	Provide an overall evaluation of the vegetative growth for Fort Riley in the year 2012.	2

Appendix D - Training Presentation Handout

Fort Riley Information

- Your role in the experiment: put yourself in the shoes of a land manager at Fort Riley.
- As a land manager, you are concerned with vegetative growth at Fort Riley.
- Factors that affect vegetative growth
    - Treatment/condition of the land
    - Time of year
    - Weather (temperature, precipitation)



NDVI Data

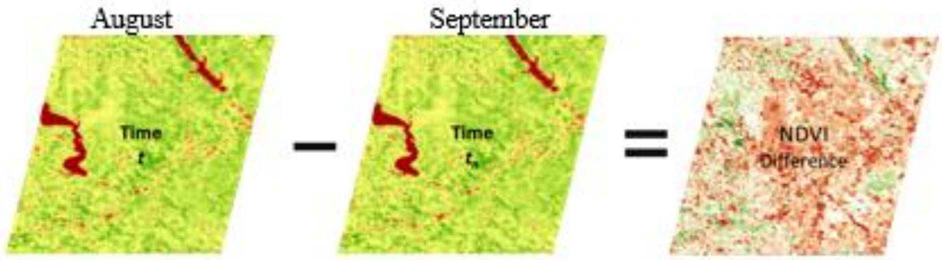
- We can use Normalized Difference Vegetation Index (NDVI) data to understand vegetative growth for a given area
- NDVI data measures the "greenness" of an area
- NDVI is calculated using satellite images and results in a numeric value for each pixel of the satellite image
- These pixel values are then assigned a color to represent greenness
    - High pixel value (250) = rich green color
    - Low pixel value (0) = dark brown color
- Images are generated over a 16 day period resulting in 23 images per year

NDVI Difference Maps

NDVI images can be used to calculate difference maps that show the percent change in NDVI values from one image to another image later in the year.

Red areas = decrease in greenness
Green areas = increase in greenness

Figure 20 Page 1 of experiment session handout.



Phenology: the study of periodic plant life cycle influenced by seasonal changes in climate.

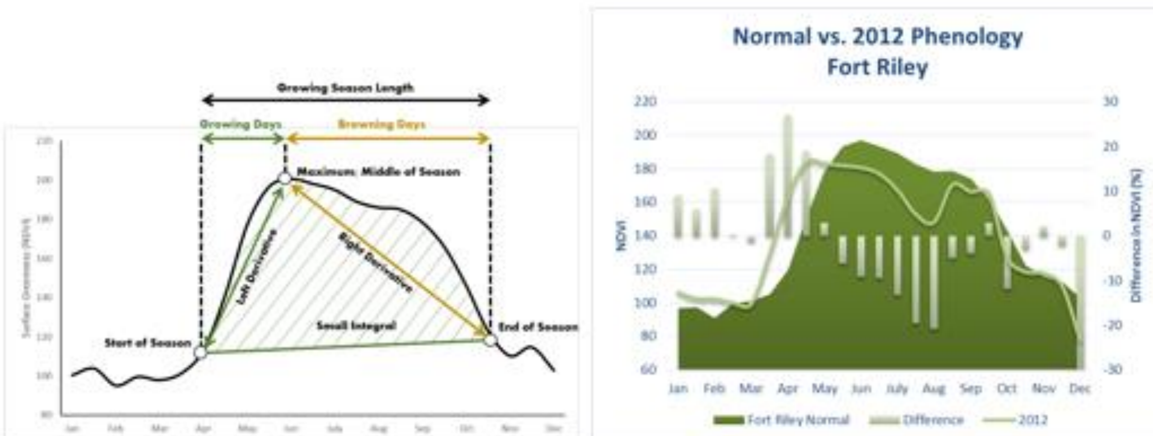


Figure 21 Page 2 of experiment session handout.