**Students' Errors Using Geographically Variable Data to Support Scientific Predictions**

**by**
**Sarah Fick**

**A thesis submitted**
**in partial fulfillment of the requirements**
**for the degree of**
**Master of Science**
**Natural Resources and Environment**
**in the University of Michigan**
**April 2012**

**Thesis Committee:**
Professor Paul Mohai (Chair)
Professor Nancy Songer

## Table of Contents

## Abstract

There is a growing need for students to be able to interpret and analyze large-scale datasets. Students have shown difficulty in interpreting large-scale climatic datasets, though it has been unclear in which tasks the students' difficulty manifests. This study looks at the types of errors made by eighth grade students in the interpretation and analysis of geographically variable average temperature and precipitation data. Four categories of errors – content errors, data practices errors, scientific practices errors, and directions errors – were identified and described. These characterizations of students' errors can be used to develop materials that scaffolds their abilities in these areas.

# Introduction

## The Problem

As our understanding of global environmental problems increases, students and citizens will need to interpret global, or regional datasets, such as continental climate data. Research has previously focused on students' use of data visualization. Graphs (i.e. Shah, 1999) and maps (i.e. Gordin, Polman, & Pea, 1994) have been explored as means of assisting students' interpretation of data. It is often assumed that these representations of data are easier for students to interpret than lists of data. Working with graphic representations allows students to work with the same analytical tools that scientists use (Gordin, Polman, & Pea, 1994). Understanding the present environmental problems to be globally based in source and impact, the scale of the problem means information is increasingly displayed in geographically variable datasets. For students to fully understand the source and effect of our environmental problems, both global and regional, students need to experience the interpretation of continental climate datasets.

Interpreting representations of global change is of particular importance to education about climate change, where students require understanding of large-scale phenomena that will influence their local environment and its biodiversity. The research of Edelson and colleagues (Edelson, Gordin, & Pea 1997) acknowledged that students have difficulty in interpreting geographically variable data. Their research focused on how the interface for accessing the data can be changed from a complex scientific tool to a student friendly analytical scaffold. They did not address the difficulties that students have in interpreting such data. This study is focused on identifying the categories of errors that students make when interpreting, analyzing, and applying geographically variable data.

# Theoretical Framework

## Scientific Practices in Learning

Within the field of science education there is consensus that argumentation is an important part of science learning (National Research Council, 2011). Argumentation is an example of a scientific practice. Scientific practices are the processes necessary for conducting a scientific experiment including making predictions, developing questions for further investigation, deciding what data to collect, modeling, collecting data, representing the data, interpreting and evaluating the data, and using the results to develop and refine explanations (Duschl, Schweingruber, & Shouse, 2006). Recent national documents have emphasized the importance of instructing students for content and practices together. The important role that scientific content knowledge plays when participating in scientific practices is the focus of a number of studies (e.g. McNeil & Krajcik, 2007; Sadler, 2004; Sandoval, 2003; Vekiri, 2002). These studies conclude that practices become very difficult for students with poor content knowledge (Vekiri, 2002; McNeil & Krajcik, 2007). Sandoval (2003) found that having students create physical links between the evidence examined and the explanations created led to more meaningful explanations. Sadler's (2004) review of studies, on the relationship between conceptual understanding and informal reasoning, concluded that most empirical studies in this area demonstrated that conceptual understanding is crucial to informal reasoning.

Together these studies show the importance of quality instruction for both the content and practices being emphasized.

When students are faced with a new representation intended to help them learn new content knowledge, generally, their understanding of the content is still in need of development and refinement. When students use a content-based representation, often their interpretation of the graphic is less than complete (Vekiri, 2002). Even when learning the content, students will not delve into details of the representation sufficiently to tell a real-world story based on the information presented (Vekiri, 2002), they should be able to glean some factual information from it. This is the basic understanding, which students should draw from to develop their own content-based conclusions.

## Predictions as a Scientific Practice

Scientific explanation is considered a high leverage scientific practice, based on its authenticity, by most national reform documents including the, American Association for the Advancement of Science's (AAAS) Project 2061 Science Education documents, the National Research Council's (NRC) Framework for K-12 Science Education (NRC, 2011) and National Science Education Standards (NSES) (NRC, 1994) and College Board's Standards for College Success (College Board, 2010). Scientific explanations give students the experience of supporting their ideas about the natural world with evidence and previously established scientific theories.

Science education models of scientific explanation, used by NSES, AAAS, and College Board, are based on Toulmin's model of argumentation (1958). These models decompose the explanation down into the parts claim, evidence, and reasoning, adapted from a larger set of categories that Toulmin (1958) used in his book *The Uses of Argument*. Claim is frequently defined as "a sentence that answers the scientific question," (Songer, Shah, & Fick, in press) evidence is "observations, data, or information that helps you answer your scientific question," (Songer et al, in press) and reasoning as "why your evidence supports your claim. Reasoning can be a scientific definition or idea to explain why you chose the evidence you did." (see Songer et al, in press) The concept of reasoning has been found difficult in both classroom and clinical studies. The construction of claim, evidence, and reasoning serves the purpose of assisting learners in construction of explanations. Researchers have developed these models of argumentation to decompose the complex practice down into more manageable parts.

This decomposition of the practice into several tasks is the initial scaffold for learners in construction of explanations. Decomposing explanations and other scientific practices breaks the process down into steps that are easier for students to complete. Several models exist for decomposing the practice to become more useful for learners (Songer et al, 2009; McNeill et al, 2006; Windschitl et al, 2008). Decomposition is one example of scaffolding of scientific practices. Scaffolds are forms of assistance that help students learn to complete a task on their own, and over time are removed as students no longer need them.

## Research on Scientific Predictions

Students' construction of predictions has traditionally worked within the model of the scientific method, focusing on the formulation of hypotheses and hypothesis testing

(Davis & Linn, 2000; White & Frederiksen, 1998). This model focuses on the ways that a hypothesis requires revision. Alternative uses of predictions have been proposed. Instead of focusing on the predictions as requiring revision, this use focuses on their value as a tool for helping students understand how their ideas have changed (Linn, 2006). There is one clear difference between a hypothesis and a prediction. Predictions require situating the statement within the context of evidence or scientific principles, while hypotheses generally do not require supporting material.

Similar to hypotheses, predictions make statements about things that will happen in the future. A major difficulty for students in their construction of predictions is in the development of coordination between existing theories with the new knowledge being generated (Kuhn, 1993). Basing a model for the construction of predictions on the model for explanations, students need to make a claim about the future (see Lee & Songer, 2003). This means that in a prediction, students might have either evidence or reasoning, or potentially both about a future time. Since a student can use either evidence or reasoning, in support of a prediction a new term is needed; in this case the term *justification* is used. For students, justifications are defined as data, a scientific concept, or a definition that supports an answer to a scientific question about the future.

Scientific predictions are similar to scientific explanations in their use of both a claim and evidence or scientific principles to support the claim (see Lee & Songer, 2003). This similarity of both form and components can be used to make an argument for the generalizability of the knowledge gained from the study of explanations to the less studied practice of prediction. It is assumed that students would struggle with the practice of scientific predictions. Students' selection of justifications to support their scientific predictions involves a similar process of selecting evidence from data, as well as potentially scientific principles. Students' difficulties with reasoning would be expected in their use of justification

## Students' Scientific Reasoning

Research has shown that students have difficulties selecting evidence from data to support their scientific explanations. Evidence in the context of a scientific explanation is often broken down into parts, in order to ensure that students provide multiple pieces of evidence to support their claims. This scaffold is intended to support students' use of sufficient evidence. Songer and colleagues (Songer, Kelcey, & Gotwals, 2009) use two pieces of evidence as their marker of sufficient evidence with upper elementary students. Students' trouble with selecting evidence from data includes using experiences with phenomena as evidence within an explanation, though those experiences are not always related to the claim. While evidence has shown fourth graders capable of selecting evidence (Songer et al., 2009), when students get to the sixth grade they still sometimes have trouble determining which data counts. The difference in abilities at the various levels is likely due to the differing complexity of the representations of the data and associated science content.

Ruiz-Primo and colleagues (Ruiz-Primo, Li, Tsai, & Scheider, 2010) broke the category of evidence down into coding questions and criteria to develop a system for establishing quality for students' explanations. For their review of students' use of evidence, the authors broke the category down into three elements: what type of information was used, what form that information took, and how many pieces of evidence were provided (Ruiz-Primo et al., 2010). They found that based on the analysis of students' explanations, only 18% of students provided complete explanations including claim, evidence and reasoning (Ruiz-Primo et al., 2010). Approximately 9% of the students provided only data, with no tie to a claim or reasoning. While 26% of students in their study were able to identify data patterns, students did not use patterns as evidence for their conclusions (Ruiz-Primo et al., 2010).

In their construction of explanations students have also been shown to have difficulty in their use of reasoning to support their claims. In their study of expert-novice differences in reasoning, Schunn and Anderson (1999) focused on the differences between domain experts, task experts, advanced beginners, and novices. They found that the novices were more likely to focus on experience based evidence or reasoning. Schunn and Anderson (1999) suggested this might be due to the novices' lack of understanding of what counts as evidence and reasoning in the field.

## Student Learning and Its Role in Prior Knowledge

This research draws from a socio-cultural view of learning (Vygotsky, 1978), which is built on the prior knowledge, skills, and experience that students bring to a learning experience as a foundation for developing their knowledge and abilities in that area. Socio-cultural learning focuses on the construction of knowledge using tools and resources to assist the student. The ideal learning experience is visualized as the individual tutoring of one student by a more knowledgeable other. In this ideal, each student gets the support s/he needs for their learning, and once the student is capable of performing a task individually they no longer receive support on that task. This ideal is impossible in a classroom environment where there are frequently twenty to thirty students in a classroom. Curriculum reform is therefore focused on how to give students the support they need to learn new content and practices at the moment they need it.

Vygotsky (1978) characterizes students' abilities on a task as having two levels: first, the unassisted level at which students can perform without the assistance of a more knowledgeable other or additional tools, and second the level of students' abilities on that same task with assistance. The difference between a student's unassisted abilities and assisted abilities make up the zone of proximal development. Students' unassisted abilities in this case characterize their capabilities for performing on a task based on their prior knowledge, skills, and experiences, and the tools available to them in the task. Implicit in this is that students will make errors on tasks that are outside of the unassisted ability level of their zone of proximal development when asked to do those tasks without assistance. When a student completes an assessment task in which some of the questions are outside of the students' unassisted ability, it is easier to identify the tasks on which students will need assistance. Through the identification of the tasks that students do and do not need assistance with, scaffolds can be targeted to the areas in need of assistance.

## Students' Prior Experience Influences How They Approach A Novel Problem

Research on prior knowledge serves as a foundation for understanding where students begin their learning experience, and what tasks those students might be able to do unassisted. It also serves as a foundation for understanding the nature of resources and tools that students need to complete the interpretations and analyses expected of them.

When students are given an assessment task without instruction related to the skills and knowledge emphasized in the task, they make errors in areas where they lack experience and prior knowledge. It is anticipated that the influence of those elements can be overcome with appropriate support and scaffolding for students. Therefore characterizing the nature of the prior experience and its influence on students' success or difficulty with the task is less important than knowing the processes that led students to make an error when completing the task.

## Students' Experience with Representations of Geographically Variable Data

Students are frequently shown representations of scientific phenomena in textbooks and science materials, and it is assumed that students will be able to interpret the representations, coming to the same conclusions that the creator of the representation intended. Research has shown that not all students look at maps showing variations in broad categories and come to the same conclusions (Bausmith & Leinhardt, 1998; Leinhardt, Stainton, & Beusmith, 1998). For example, a dataset of annual average temperatures and precipitation values could be presented in five degree Celsius bands of temperature and 50cm bands of precipitation. In order to understand how the temperature and precipitation will change as a result of climate change, students must understand what climate is, and how it is represented in these data layers. Their ability to interpret and analyze the information presented in these layers has implications about their ability to explain these changes and support those explanations with evidence.

## Research Question

This study is designed to inform our understanding of the research questions: What types of errors do eighth grade students make when interpreting and analyzing geographically variable climate data for use as justification for a scientific prediction? How do the types of errors that students make relate to the scientific practices being used for interpreting and analyzing geographically variable climate data for use as justification for a scientific prediction?

## Research Methods

An assessment was developed that required students to use three geographic interpretation and analysis skills (identification, pattern identification, and data manipulation) to answer questions associated with geographically variable data. The questions required a variety of levels of prior knowledge of scientific and geographic processes in combination with interpretation and analysis skills. This combination of content knowledge and skills was meant to be a realistic replication of a classroom task.

### Design of the Assessment Task

The socio-cultural learning theory assumes that students' prior knowledge and experiences will lead them to interpret and analyze geographically variable data in different ways, and will lead them to choose different evidence. Each of these factors will vary by student, giving each student a different experience with the task. Since the task is not designed to measure student learning, there is no need to control for these elements, and the ways that they influence a students' success is less important than the students' responses. This assessment is intended to demonstrate the difficulties that students have with the task, resulting in errors.

Middle school students have had some experiences with maps in a variety of contexts and circumstances both in and out of the school environment. These experiences serve as a foundation of knowledge and skill that students bring to the analysis of spatial data, which requires some of the same skills involved in interpreting a graph. This knowledge that students bring to the learning task is the students' prior knowledge relative to the skills required in the task. The goal is to characterize the range of abilities through identifying the questions with which students struggle. The assessment task was designed so that the knowledge and skills as well as the errors that students made while interpreting and analyzing geographically variable data in the context of science content and practices could be highlighted. Students used average temperature and average precipitation data (figure 1) to develop scientific predictions. In designing the assessment task, several constructs were identified as being necessary for interpreting and analyzing geographic data for the development of scientific predictions: identification, pattern recognition, data adjustment, describing the cause of a pattern, and making predictions including a claim and justifications. These constructs are in line with those skills expected of middle school students by the national geographic standards (Bednarz et al., 1994).

## Data Analysis Constructs

The assessment was designed to measure students' abilities to interpret, analyze, and apply geographically variable data in content dependent and independent situations. Very specific geographic interpretation and analysis skills were used as measures of interpreting, analyzing, and applying the data. These terms are outlined below with examples provided to illustrate the constructs. Each of the activities in the assessment was focused on a particular practice, and also required some content knowledge, which was not instructed for before the assessment.

The term "interpretation" represents the process of finding the locations of a value, or range of values, found on a map. Doing these skills requires the ability to do three things: recognize which of the data ranges on the legend represents a given range of values, identify the color aligned with that range of values, and find major patches of locations represented with that color on the map. Activities that required interpretation were considered to be practice focused.

The term "analyze" includes two processes: pattern recognition, and the constant modification of data values. The first is the ability to recognize patterns in geographically variable data. In the assessment, students were given the definition of a pattern as: "When something is placed in a way that is not completely random [statistically random]. There is an order to the way things look." Pattern recognition requires that students notice when there is a non-random organization of the data. The recognition of patterns requires the ability to observe regularity. When students are asked to describe a pattern, that process requires some geographic or scientific prior knowledge. Describing a pattern or explaining why it might occur requires knowledge of geography and scientific processes that might cause the pattern. Pattern recognition was considered a content dependent scientific practice. While students might observe a pattern it is possible they do not recognize it as such unless they have supporting scientific content.

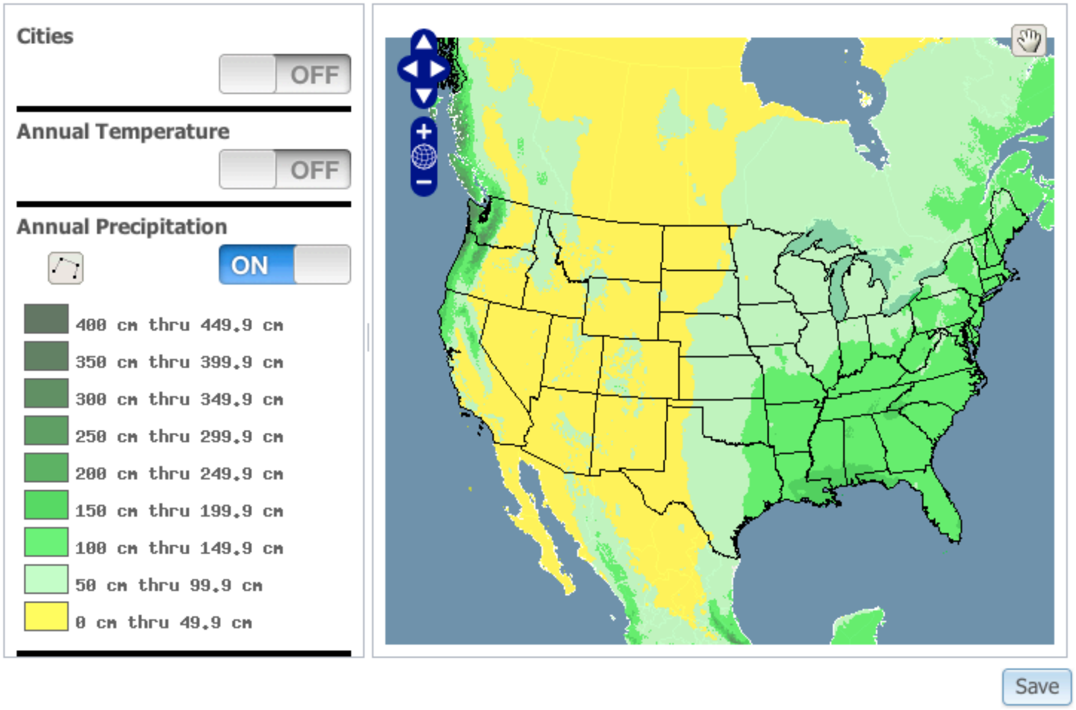
The second type of analysis task is the adjustment of data values by a constant change to represent changing conditions. After recognizing magnitude and direction of the change, students then adjust the values represented by the various colors to reflect the change, and circle the new areas satisfying the data range. Data adjustment was considered a scientific practice, because it requires students to visualize a constant change in the data.

Finally, the students must apply observations from a map as evidence to support an answer to a scientific question. This requires the student to select a supporting observation or pattern from the data represented on the map associated with the answer to a scientific question and to describe that observation or pattern in words as evidence. Predictions include a claim, and two justifications. The two justifications can include data, a scientific concept, or a definition that supports their answer to the scientific question. Because this task was intended to focus on interpretation and analysis, students' knowledge of scientific phenomena and political labels was intentionally kept to a minimum in the assessment task.

This was done by accepting any accurate characterization of the students response. For example, if a student said, the area that is blue and yellow, and it accurately described the answer that was acceptable. Students did not need to know the names of states, but were expected to know where the border of the United States was, since questions required students to "circle the areas within the United States". The content was limited in this task, but content was still necessary. It was considered a content dependent scientific practice.

## Study Population

The study consisted of nine eighth grade students in a science focused charter school in the urban center of a large mid-western city. The school serves students living anywhere in the city district, which encompasses the entire city. The eighth grade class that year was composed of 96.4% African-American students, the remaining <4% being made up of other races; 61.21% of the students were male, with only 38.79% female. Less than 10 students in the school were classified as English Language Learners. The school was a charter school in a very low SES school district. Of the nine students interviewed, two students were female, and one student was of an ethnicity other than African-American. The students were from different classroom sections, but had the same science teacher. The teacher selected students who represented a range of abilities, exhibited by their participation in science class to participate in the study. Since the participants were at the very end of their eighth grade year, they can be considered equivalent to students transitioning from middle school to high school students. They had attended the charter school for one to three years. Students who had been there for two or three years had experience with a research-based NSF sponsored curriculum focused on instruction in both scientific content and scientific practices. The scientific practices emphasized in this curriculum were data collection, analysis, and explanation. Students with only one year's attendance were not exposed to this curriculum, but may have had experience with some of the scientific practices in coursework during the academic year.



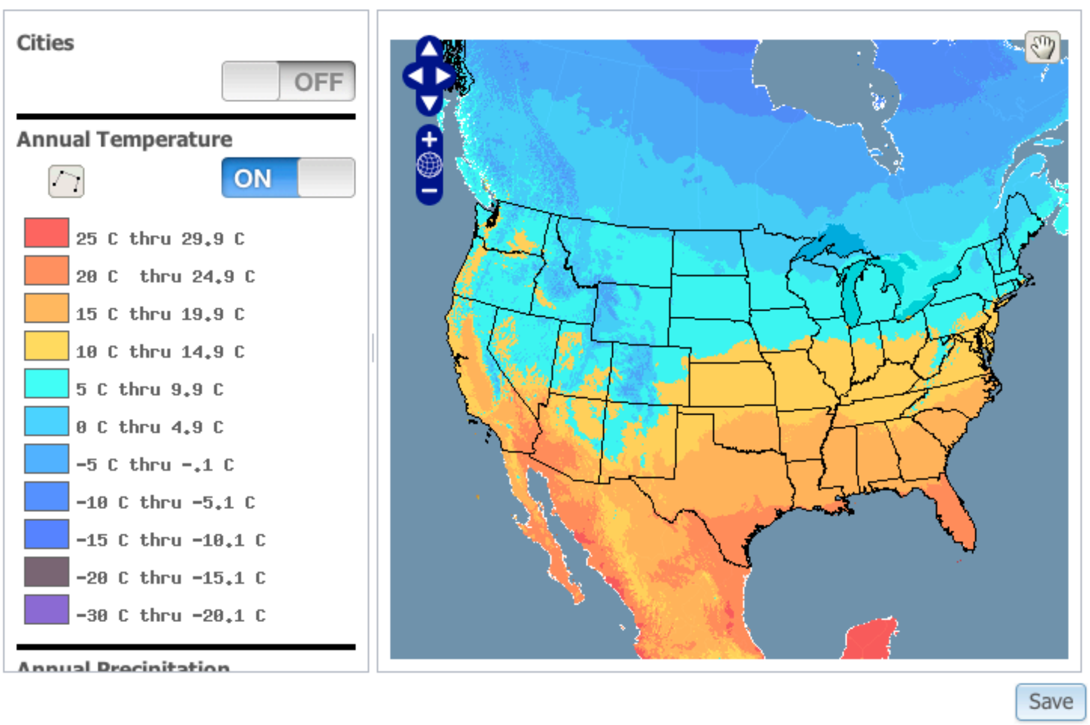




Figure 1. Maps of average annual precipitation (top) and temperature (bottom) classified by 5 degrees Celsius and 50cm of precipitation, which are broad classifications of data.

All of the students intended to attend public or charter high schools for ninth grade. The student participants interviewed would be considered minority students based on their race or ethnicity.

## Data Collection

Students completed the assessment in their science classroom with the teacher present, but in a different part of the room. The teacher was not participating in the assessment process or paying attention to the responses. There were other students present in the room, at times, working on other assessment tasks. The school was not a quiet work environment for many of the participants as there was a lot of end of the year excitement in the hallways. Participants, working individually, were presented with an online interface containing a combination of maps and questions on each page. Each participant was given the 19-question assessment (See Appendix A for the full assessment), which took from 40 to 70 minutes to complete. All questions focused on the interpretation and analysis of geographically variable data. Tasks were designed to be dependent on varying amounts of geography and science content knowledge. The questions varied in difficulty and amount of required content knowledge. The required geography knowledge was basic North American political boundaries, while the science knowledge focused on continental level climate trends and basic habitat knowledge.

Students were provided assistance with use of the interface to prevent usability from being a limiting factor in their success. The author was present as students responded to the questions. Using a think aloud protocol (Ericsson & Simon, 1993), the author prompted the students to provide information about the process they were using to answer the questions, and why the students selected the particular answers.

Using ScreenFlow (Telestream, 2011) software, students' interpretation and analysis processes were recorded. ScreenFlow software stores a continuous display of the screen of the computer with coordinated sound of the user speaking. Video of the user can also be taken for computers with a camera. In this circumstance only the screenshot and user's voice were recorded. Recordings consisted of: a continuous screenshot of the students' screen as they completed the activities accompanied by audio of the students' described responses, the think aloud description of the process they used for answering the question, and the written responses students provided for the assessment questions.

## Data Analysis

Students' responses consisted of: a continuous screenshot of the students' screen as they completed the activities accompanied by audio of the students' described responses, the think aloud description of the process they used for answering the question, and the written responses students provided for the assessment questions. Iterative rounds of coding (Miles & Huberman, 1994) initially characterized written responses as correct or incorrect. The coding was based on a pre-established notion of what was correct and incorrect using the rubric of correct responses. Subsequent coding used incorrect responses as the basis for the development of codes descriptive of the type of error made, following the model of grounded theory based coding (Patton, 2002). Codes were applied on the response level, where each student provided two or three types of response to each question: verbal, written, and when required by the question a circled map.

Students' responses varied from circled areas on a North American map to verbal description of their responses, and verbal descriptions of the processes they used to answer the questions. The students' verbal and written responses were transcribed. This process used both the audio recording and screenshots associated with the ScreenFlow (Telestream, 2011) recording of students' progress. These verbal descriptions and responses were transcribed. The researcher transcribed the students' circled responses by creating a written description of the circled area. The description highlighted the difference between the correct response for that question and the area the student circled.

## Round I of Coding

Both students' verbal responses and their written responses were coded as correct or incorrect as compared with a rubric established for the assessment questions. Responses coded as "correct" were not further examined.

## Round II of Coding

Two assessments were coded with codes that were descriptive of the errors the students made in the incorrect verbal and written responses. Separate codes were used for the verbal and written responses, giving each question two sets of codes. Additionally, each question was coded with codes to reflect all the ways that the responses were observed to be incorrect in the initial round of coding including: word choice, type of information provided, content errors, and practice errors. After two assessments were coded the codes used and a description of what errors the codes represent were elaborated in a codebook. The codes created in the initial two assessments were used to code the remaining assessments. Novel errors encountered resulted in new codes. In an attempt to ensure that all errors students made were coded, each response was given as many error codes as corresponded to the errors, thereby ensuring that all errors that students made were represented. This resulted in many responses with multiple error codes, some individual responses having as many as six codes. The codes were grouped into broad categories. These categories, "bordering errors-major, bordering errors-minor, geography errors, science errors, description errors, type of information errors, and consistency errors" (described in the subsequent section) indicate difficulties with particular elements of the process. These codes were condensed to reflect the construct that was associated with the student's error. The specific error codes were elaborated in a codebook describing the meaning of the codes being used. To gain a fuller understanding of why the student made the error in question, students' responses were condensed so that all of the students' responses to a particular question were used when coding that question, thereby reducing the number of codes per student response from two or three to one code.

## Round III of Coding

Subsequent rounds of coding focused on the part of the activity that resulted in the student making an error. This decision was based on further analysis that determined that some codes were hierarchical. For example, students could not circle the wrong area on the map if they did not understand the directions. Therefore errors associated with the task directions highest level. It was also determined that the student's description of the process they used, and their verbal and written responses should be used in combination to confirm or disconfirm initial ideas of which category of error the response belonged in. Using all three elements to determine the code allowed for a fuller picture of the process when classifying students' responses.

Responses that had been given multiple codes were reviewed to determine which error was the dominant error, and assigned the corresponding error code. This decision making process was based on a greater reliance on the context of the responses. The decision for the error code was based not only on the students' written response, but also their verbal response and the students' description of the process they used to solve the problem. This reduced the number of coded elements from three to one, and also increased the grain size of the analysis. This analysis included the information gained from all three elements in the code assigned to the inclusive student responses. Error codes were regrouped to represent similarities in the types of knowledge or practices that students had difficulties with. The resulting categories can be illustrated by the specific errors that fit within them (figure 2). Diagrams organizing the flow of errors and the types of errors throughout the paper were made with Inspiration (Inspiration, 2011) software.

## Results

The coding process revealed categories of student errors. These errors and their relationships to sub-categories of errors are shown in figure 2. The figure should be read as a flow chart, with errors on different levels representing a hierarchy, and codes on the same level, being mutually exclusive but not hierarchical. In figure 2, the major categories of error are represented as orange circles, the highest level of code for incorrect responses, and subdivided into the various means of making the error. Examples of the semi-major errors are represented by yellow rectangles. The blue rectangles represent the specific errors. Those rectangles with blue hashing and a magenta border represent examples with more than five observations, qualifying those examples for

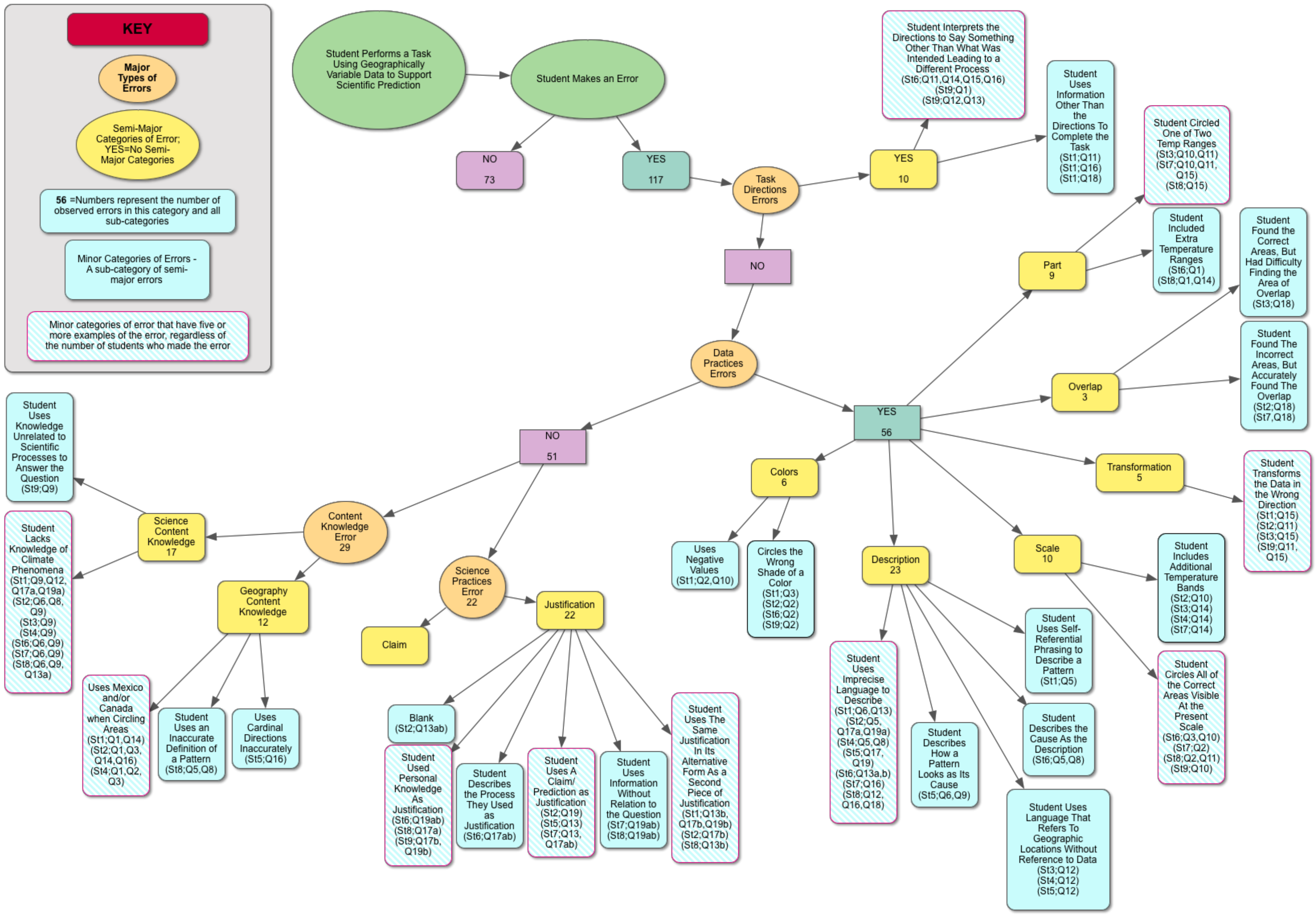



Figure 2. A flow chart of the errors that students made on the assessment task. Each orange oval is a major type of error. Each yellow rectangle is a semi-major category of error. The examples of the errors are in blue rectangles. The "S" number represents the participant who made the error, the numbers following the "Q"s correspond to the question numbers for the assessment in Appendix A.

explication in a subsequent section. At each of the decision points where errors divide into sub-categories of errors the number of observations is written in that box. The letters and numbers in the lowest categories of errors represent the student number and question number that was coded with that error. These specific observations allow readers to observe for themselves any underlying patterns between the types of error and the questions they were made on.

## Types of Errors

The errors that students made when interpreting and analyzing fall into four major categories of errors (figure 3). Each construct was associated with one or more type of error. The major types of error are described briefly below, and elaborated with examples from students' responses in the subsequent sections of the paper.

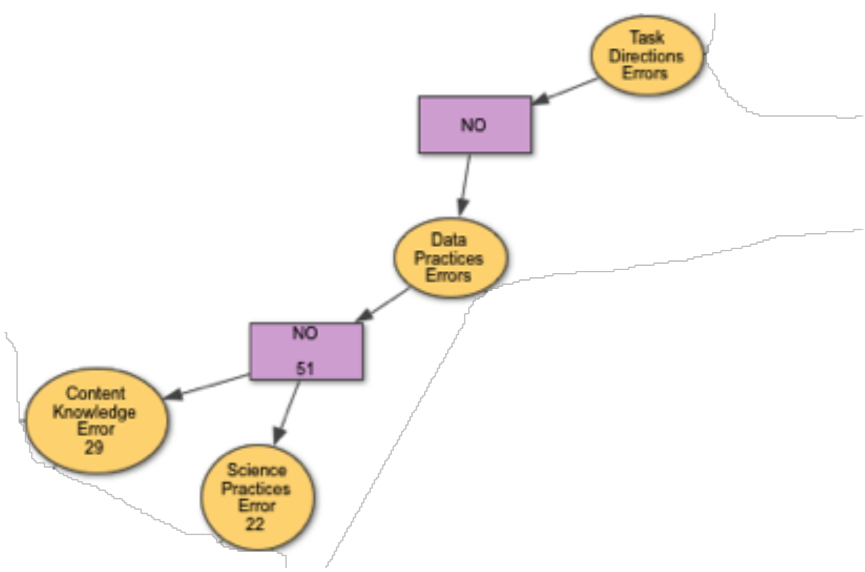



Figure 3. The major categories errors, shown in orange circles here as taken from figure 2. The major categories of error corresponded with either a type of content knowledge or a scientific practice, the exception being task directions errors, which were more broadly text interpretation errors.

## Task Directions Errors

Task directions errors were any errors that were caused by the student interpreting the directions to say something other than what was intended, leading the student to a different process for answering the question. This occurred both when students used alternate procedures and when they used information other than what was required to answer a question. Task directions errors were generally independent of students' content knowledge, but often related to the students' understanding of the practice. Being unfamiliar with a practice led students to interpret the directions in a way other than what was intended.

## Data Practices Errors

Data practices errors are associated with interpreting, analyzing and describing the data represented in the maps. Six categories of errors fell under the data practices category: circling part of the required ranges, errors in transforming the data, errors in finding areas of overlap, errors caused by using a limited scale of the data, errors in describing what they observed in the data, and errors in interpreting the colors. Data practices errors were errors in which the students were unfamiliar with a particular data interpretation or analysis process, and were based on their familiarity with the practice of working with data.

## Science Practices Errors

Science practices errors were those in which students made an error in the construction or description of their justifications. Since errors in claim were only due to basing the claim on prior mistakes, the category of "claim errors" was not used. Justification errors were due to a student using something other than related evidence to support their claim, or using less than two pieces of evidence (as required). Errors ranged from using the same piece of evidence twice to using personal knowledge as justification.

### Content Knowledge Errors

Science content knowledge errors consisted of two types of errors. The first type of error occurred when students used knowledge unrelated to the scientific process to answer the question. The second type of error occurred when students lacked the knowledge of climate phenomena necessary to answer the question.

Geography content knowledge errors consisted of three types of errors. The first type of error was when, in a question that specified areas within the United States, the students' circles included areas in Mexico and/or Canada. The second type of error occurred when students did not understand the definition of a geographic pattern. The third type of error occurred when students used cardinal directions inaccurately in their responses.

## Examples of Students' Errors With Interpretive Analysis

The errors from the major categories described above will be expanded on. The sub-categories of error that had more than five examples, regardless of the number of students who made the error, are expanded on in the following sections. Examples of the errors are presented and discussed in terms of what qualified the error for that type, and what this type might mean in terms of the process that caused the error.

Only those with five or more examples are highlighted in the following section because five examples provides sufficient evidence to compare and contrast examples, while providing plenty of detail for the case to be sufficiently described. The choice of the cases is not to place emphasis on the importance of one versus another type of error, but merely to highlight the errors that were found to be problems within this sample. A sample of different individuals would likely result in additional errors. Within these students, the following errors were made in five or more questions indicating a problem that needs attention.

## Task Directions Errors

The first place in the students' process that might have caused them to make an error in responding to the assessment question was in their interpretation of the task directions. There were two types of errors that students could make while interpreting the directions (figure 4). They could either interpret the directions as something other than what was intended, which is outlined in the discussion below, or use other information to complete the task. Only one student fell into the second category, with fewer than five examples. This discussion focuses on the first type of error.

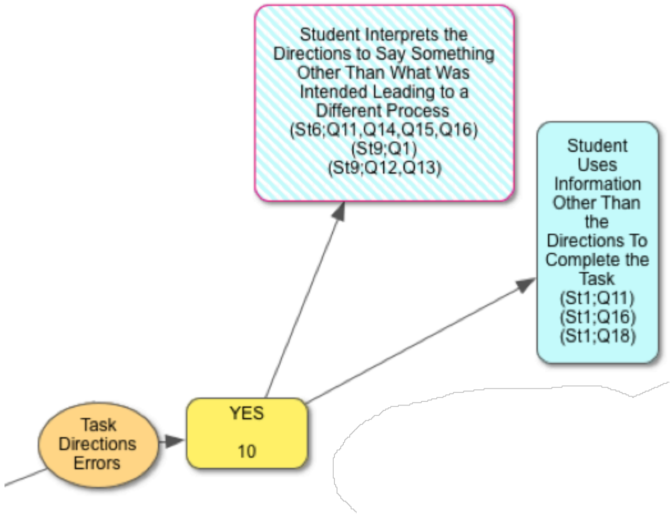



Figure 4. The branch of the complete analysis, figure 2, focused on task directions errors. Two types of errors were observed, only one of which had five or more observations (shown in blue stripes and magenta outline).

### *Procedural*

Student interprets the directions to say something other than what was intended, leading to a different process.

Even though the directions were tested with a number of students before the study group worked with them, students had a variety of ways of interpreting them, which led to inconsistent performance on similar tasks within a test (figure 4). For example, the following image (figure 5) taken from a students' response, highlights that the direction calling for a student "to border the hottest areas in the United States" (Question 1, Appendix A) can be interpreted in a variety of ways.

In this response (figure 5), there are several things to note. First, the student began bordering by using only existing borders. When the student got to California, and realized the need to include a significant portion of the state not satisfying the requirements for "hottest", the student asked if cutting through the middle of the state was allowed. This began their deviation from using only state borders. Second, the student circled more than one data range. This was particularly interesting because it meant that the student was using something other than the data ranges to determine the "hottest" locations. The student selected an arbitrary cut-off point based either on colors or on values. The student chose an atypical means of bordering the areas, which was abandoned on subsequent questions. The student's primary error concerned their interpretation of "hottest." They included a wide range of values to represent hottest, but carefully excluded areas of blue. This type of error is most easily associated with a misunderstanding of the directions. The student did not understand that only the hottest value range should be bordered, and that meant creating their own borders. In subsequent responses, the student bordered using the expected method.

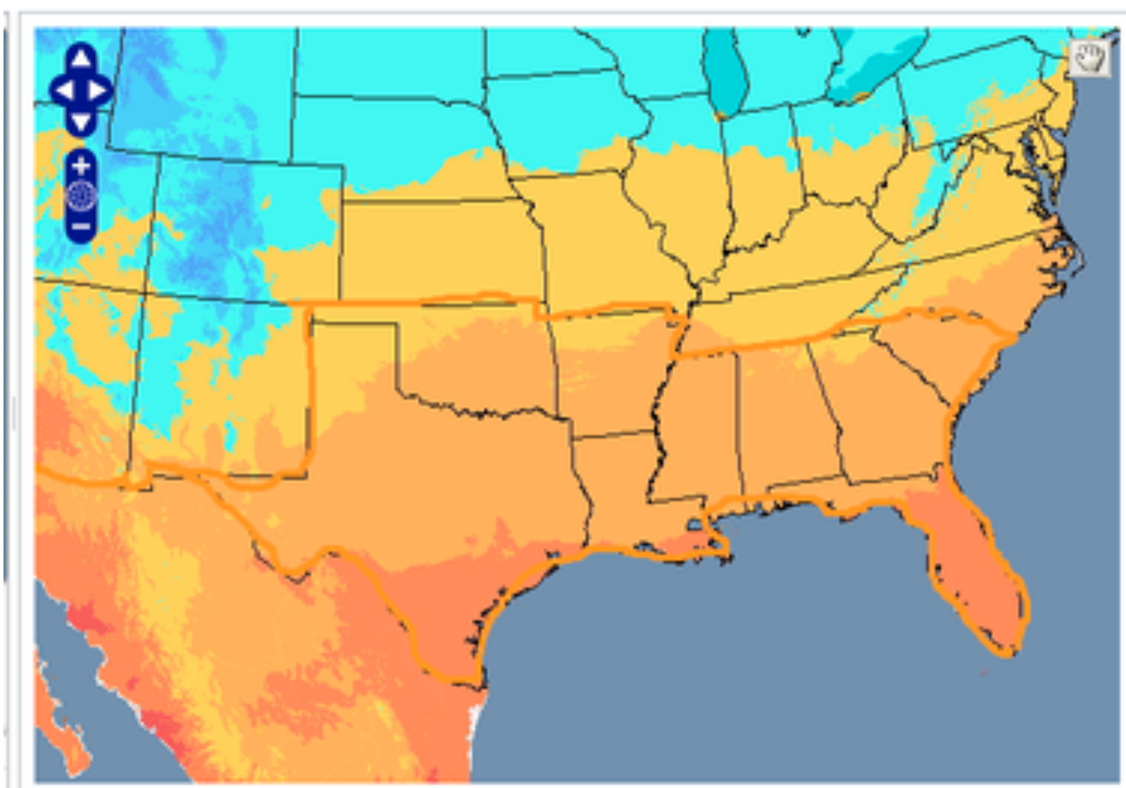


Figure 5. A students' response to question 1 of the assessment. The students' response is seen as an orange border in the orange area of the temperature map.

In a different example, "circle the area where you would expect to find the Bog Lemming in the United States, on the map above," (Question 14, see Appendix A), a student refused to answer some of the assessment questions, because the student believed s/he had already answered the questions previously.

> *"I already circled it. That's kinda like… that's the same question…. Even if I won't be able to see it really… And I've already circled it on the other map."*
>
> *– Student 6*

The student had in fact answered similar questions, but the new questions drew on a different set of data: a precipitation rather than a temperature map. This error took root in a previous error related to task directions, where the student used the wrong layers to answer a previous question. The student did not read all the directions (evidenced by the fact that the student read aloud everything s/he read), and as a result answered the above question using too much data. The student was using both the temperature and precipitation layer when the directions called for the student to use only the precipitation layer. The question asked the student to, "Circle the areas where you expect to find the Southern Bog Lemming if the United States were getting drier, 50 centimeters less precipitation everywhere." (Question 15, see Appendix A)

> *Student 6: It says, "Circle the areas of the United States where you would expect to find the southern bog lemming if the earth was getting drier"*
>
> *S6: [turns on the temperature and precipitation layers]*
>
> *S6: "I'd wanna say right there…"*

*S6: [circles the majority of MS, AL, GA, and TN on the combined map]*

(see figure 6, top map)

*Researcher: Okay, what made you say that?*

*S6: [turns off the temperature layer and starts circling the same area on the precipitation map]* (see figure 7, bottom map)

*S6: "I choose it because that was the one that made the most odd color. So it kind of stood out to me the most."*

In contrast with the previous example of task direction errors, this student's errors were caused by lack of careful attention to the directions. This chain of errors led to a total of four directions errors.

These errors illustrate two points. First, the tasks students were completing were not always intuitive. Students who did not carefully read the directions (as in the second example) were not able to complete the task. Second, the wording of tasks was extremely important. While this seems intuitive, it stresses the importance of students being able to ask for clarification during a task. This is not always the case in modern classroom assessment. The student, in the first example, read the word "border" as something he had to use when drawing his lines, thus reading more into the question than required. While students often read more or less of the instructions than necessary, it suggests that an important area of focus is the need to help students find the key elements of the directions. Both students focused on discrete portions of the directions, changing the way they understood the task as a whole.

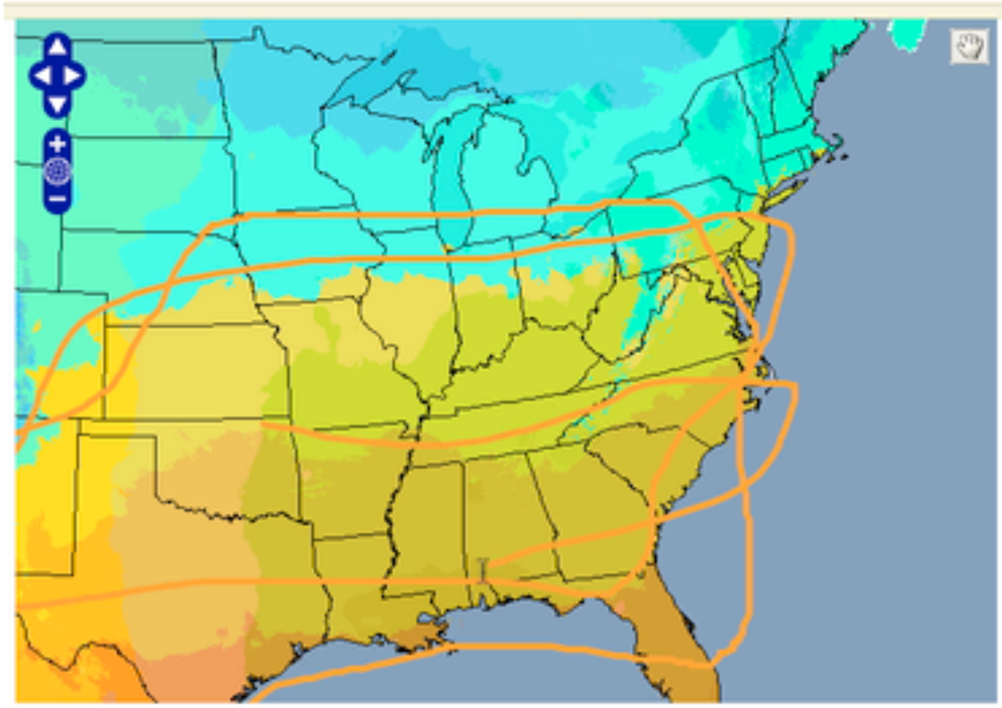
### ***Data Practices Errors***

Data practices errors were those errors in which students either incorrectly interpreted or analyzed the geographically variable data. These errors fell into one of six categories. The categories had varying numbers of sub-types of errors from one to five. Those sub-types with five or more examples are described in the analysis below. The sub-types of errors are also shown below (figure 8), the types with sufficient examples to be elaborated are striped with magenta outlines. Each of these types of error with five or more examples are elaborated in the following sections.

### Selected a Portion of the Correct Answer

*Student circled one of two required temperature ranges.*

The correct response for question 10, "Draw a border around the areas where you would expect to find the Bog Lemming in the United States, on the map above," (temperature preferences were provided, see Appendix A) required the student to circle two temperature ranges, the gold colored and the light orange ranges. The student circled only the gold colored range (figure 9), half of the required information. As a result the same student only circled one data range in the second part of the question, which required the adjustment of the circled areas by five degrees. The correct answer, for the second part, would have been to circle the light blue and gold areas. Note in figure 9, the student's response to the question is confusing considering that the student provided the wrong answer. Figure 9, shows the student's response to the first and second part of the question. Since both parts of the question are shown, it can be incorrectly interpreted as a correct answer. The student's response to part two was the light blue area only.



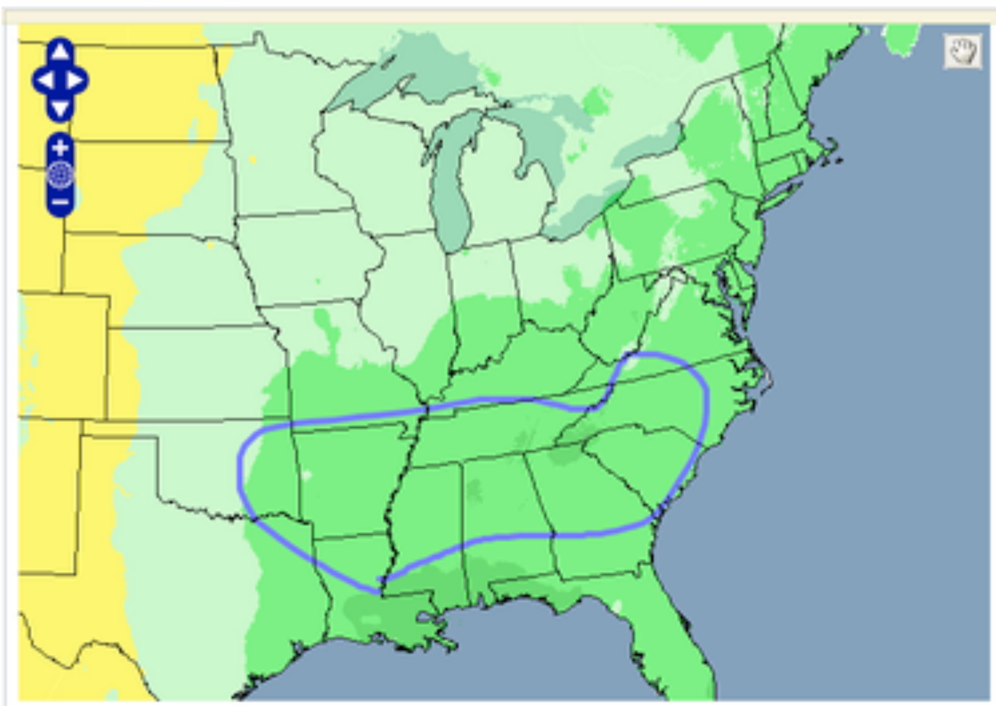


Figure 6 & 7. Students' response to the identification of a particular area on the map, both of these circled areas represent the students' response to the location where a species would live if the world were getting drier. The top map (figure 6) shows a students' circled area on top of both the temperature and precipitation layers. The bottom map (figure 7) shows the same students' circle on a precipitation data layer.

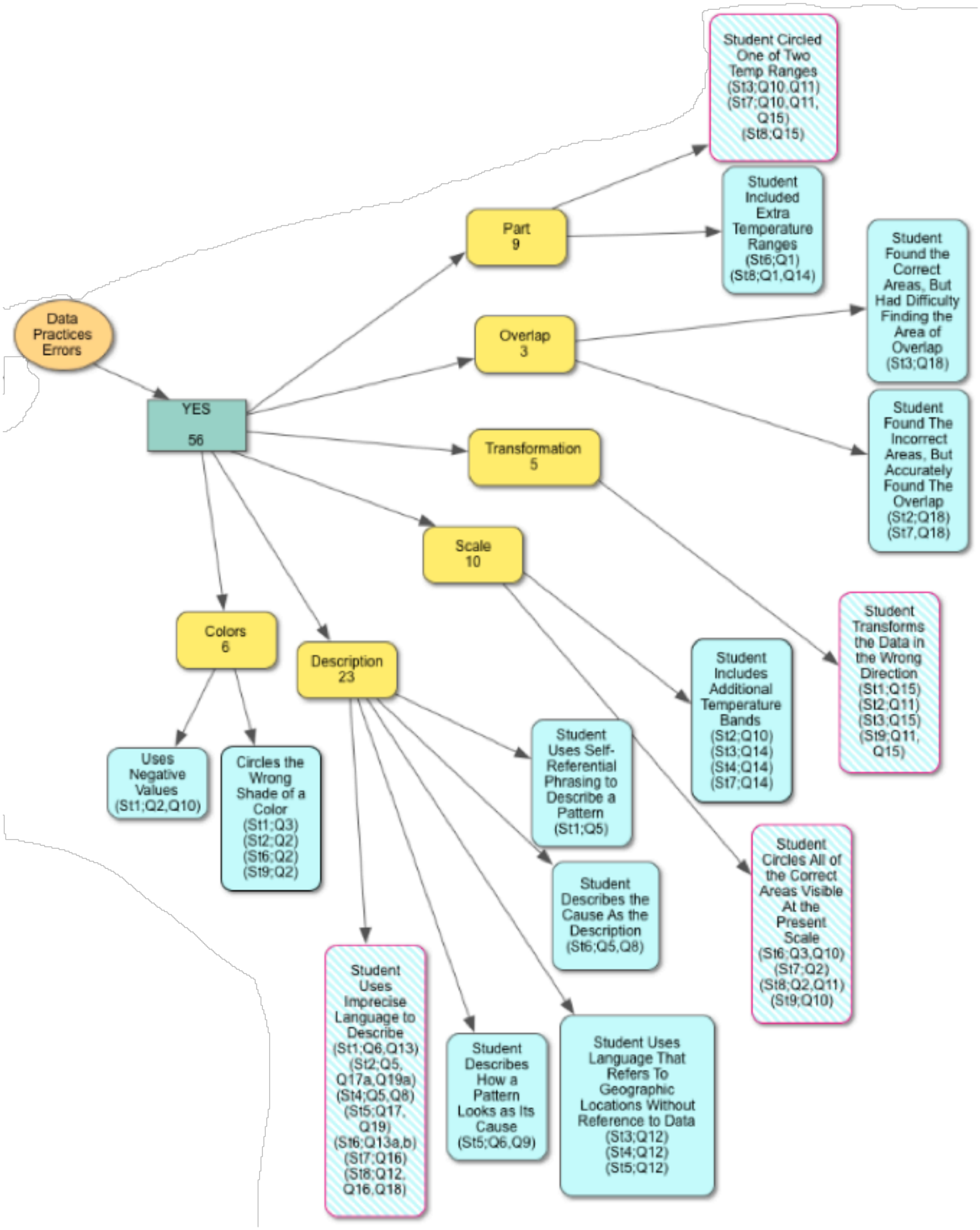



Figure 8. The branch of the complete analysis, figure 2, focused on data practices errors. Six types of practices were identified each of which had associated types of errors, the types of error with five or more observations are shown in blue stripes and magenta outline.

Because the student initially circled only one of the two temperature ranges (figure 9), their transformation only included one circled range (figure 10), and the initial circle remains). For two students, the same questions resulted in the same error, circling only a portion of the expected temperature range. Due to the structuring of the questions, if a student incorrectly answered question 10, “draw a border around the areas where you would expect to find the Bog Lemming in the United States, on the map above,” they were more likely to incorrectly answer question 11, “draw a border around the areas where you expect to find the Southern Bog Lemming if the United States were getting hotter, 5°C warmer everywhere,” (see Appendix A). This occurred even when they used the correct process for question number 11. In the example provided (figures 9 and 10), the student circled one of two required temperature ranges in question 10 (figure 9), and therefore circled only one of two required temperature ranges in question 11 (figure 10). The student did correctly transform the data previously circled, but because only one of the two data ranges was circled, they made the same mistake in question 11.This student may have lost track of the extent of the data that needed to be circled.

## Transformation Based

*Student transformed the data in the wrong direction.*

The adjustment the student made in figures 11 and 12 are examples of transformations. These transformations involve adjusting the data to determine the location of various data values provided there was a constant change in values. In the question, students were asked to “circle the areas where you expect to find the Southern Bog Lemming if the United States were getting hotter, 5°C warmer everywhere.” All temperature values were increased 5 degrees, the same range of values as in one temperature band. The task should have resulted in the adjustment of the data by one temperature range.

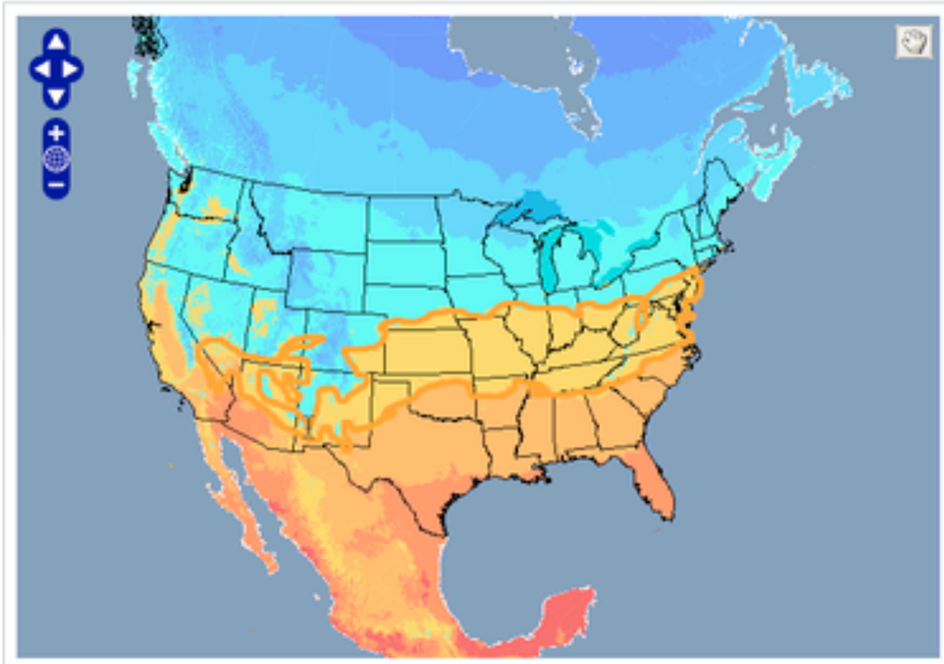
In both examples, figures 11 and 12, the student transformed the data in the wrong direction. The question requires the student to consider the location where the animal might live if the world were to get drier (in the top example) or hotter (in the bottom example). In the top example, the student circled a location drier than the animal's preferred habitat, rather than where the animal would live if the region were drier. In the bottom example, the student circled the hottest area on the map. It is likely both of these students transformed the animals' preferences instead of the data, which led to a transformation in the wrong direction.

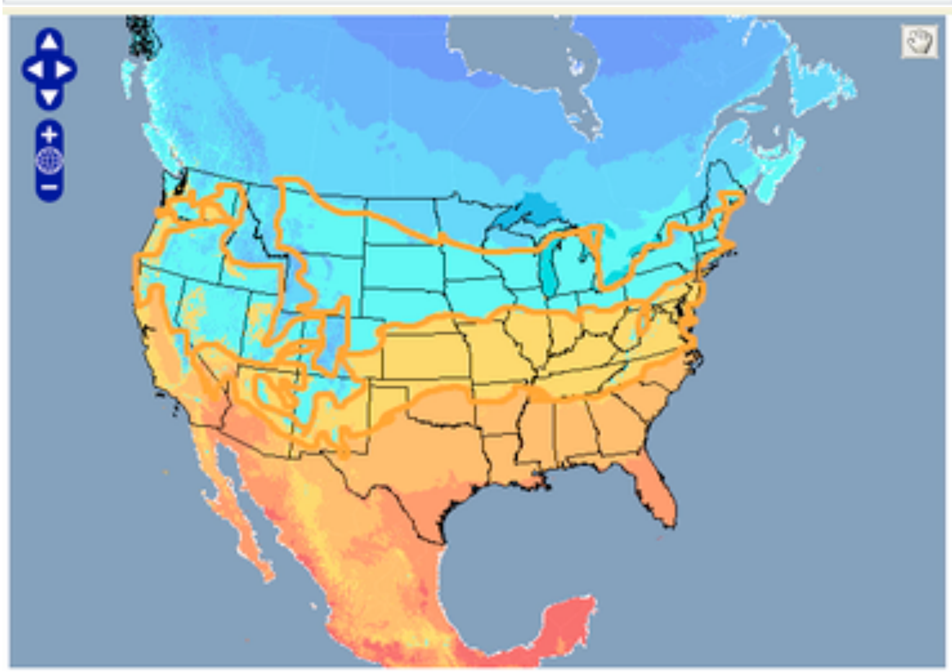
## Scale

*Student circled all of the correct areas visible at the present scale.*

The tasks requiring students to circle data demanded attention to the detail in the data, and required them to consider the data at multiple scales. This is illustrated in the mistakes of the two examples below, figures 13 and 14. In the top example, the student circled only the largest area that satisfied the requirement for the range of values.

In the bottom example, the student zoomed in to look at the detail of the data, but in doing so, missed critical areas that satisfied the required characteristics in other areas of the map. Two tasks required in circling the correct value ranges are to examine the map from a distance, to get a sense of all the possible locations, and to zoom in, to see where the smaller areas of values lie.





Figures 9 & 10. The student selected only the gold area, when the question asked the student to circle the gold and orange areas (top, figure 9). This resulted in the student only circling half of the transformed values as well (bottom, figure 10). The circle from figure 9 remains in the illustration for figure 10. The new circle is the difference between figure 9 and 10.

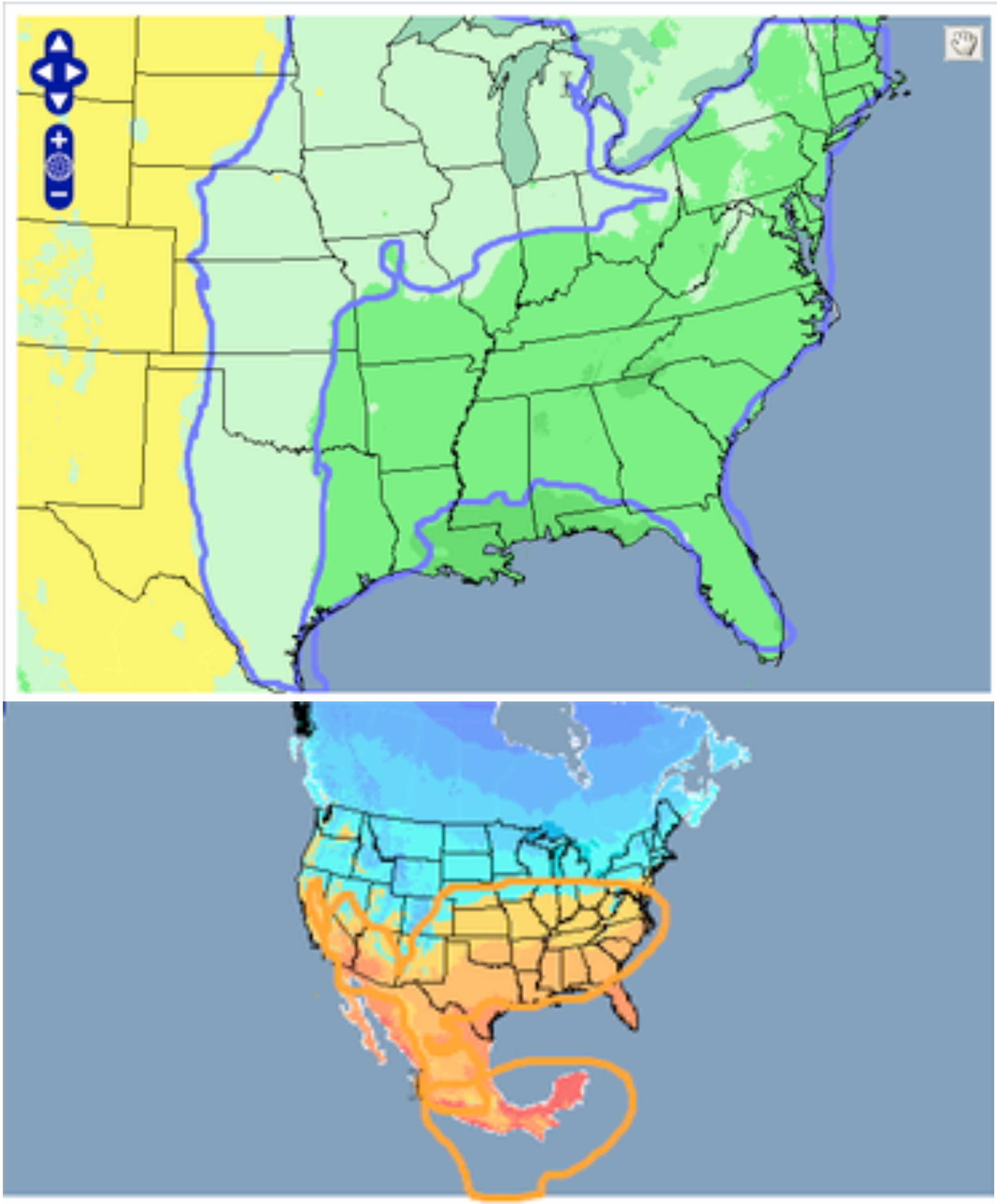


Figure 11 & 12. In the above examples the student transformed the data in the wrong direction. In figure 10 (top), the student was asked to circle where the species would want to live if the world were getting drier. In figure 11 (bottom), the student was asked to circle where the species would want to live if the world were getting warmer. In both cases the students circled the drier and warmer, locations, rather than the wetter or colder locations.

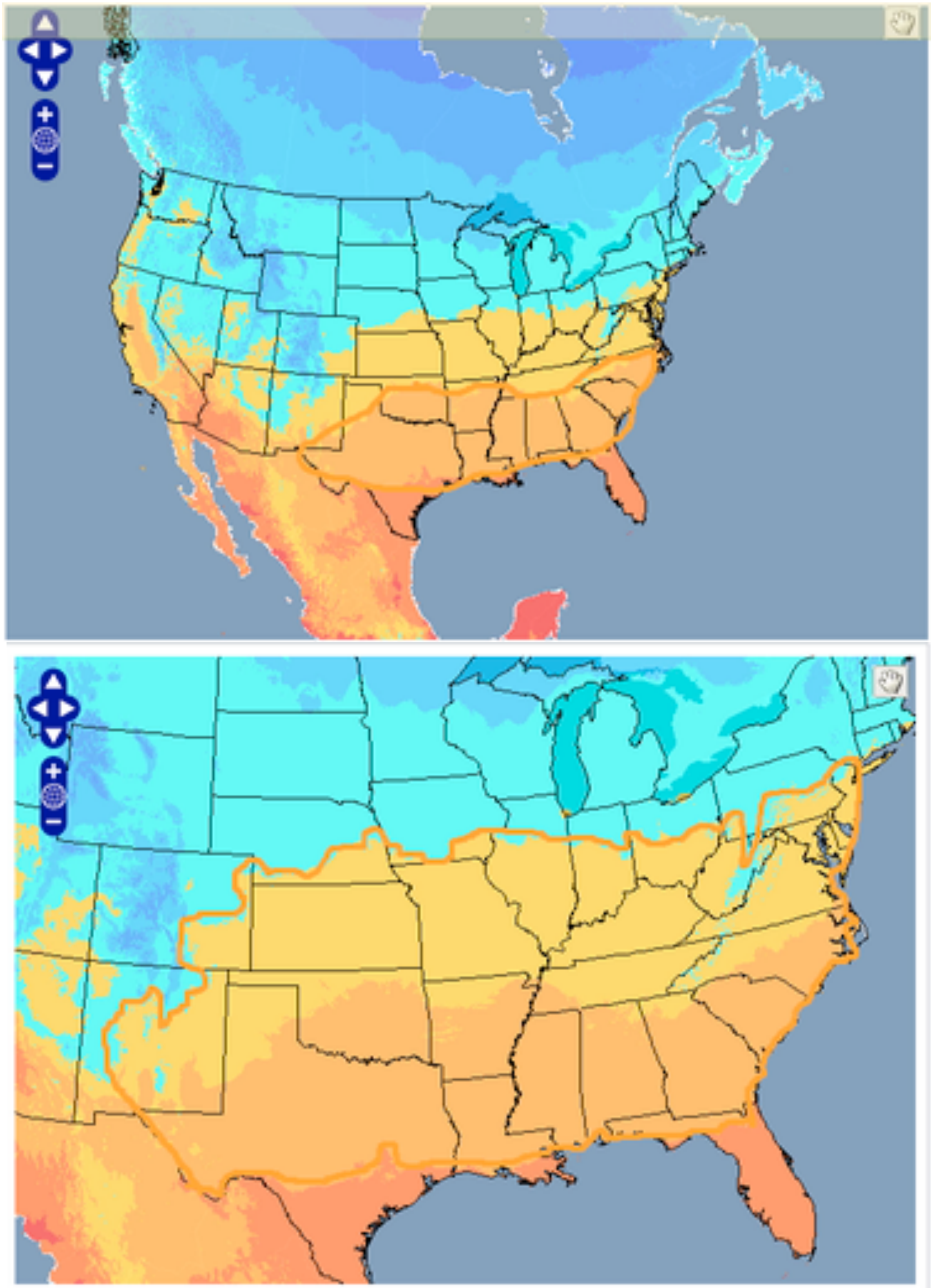


Figure 13 & 14. In figure 13 (top), the student circled only the largest area of a particular color. In figure 14 (bottom), the student circled all the areas in the visible area, but would have found more had they zoomed out.

Description Based

*Student used imprecise language to describe the location.*

Several questions in the task required students to describe the areas they identified on the map. The descriptions served to identify exactly what the student was looking at. For example, in two questions, students were asked to describe the pattern identified in the data. Students were also asked to describe the evidence used to support their claim. Students described a circled area, which served as their claim. In these questions, information taken from the map served different roles simultaneously, as a claim and as justification for that claim.

In descriptions of the patterns they observed, "in a complete sentence, describe how the pattern looks," (see questions 5 and 8, Appendix A), students frequently referred to the colors observed in the maps. Few students referred to locations or made relative statements to describe regularity in the pattern. Students only referred to the order of the colors as representative of a pattern:

> *"The pattern looks like its going from blue orange blue orange" – Student 2 (Question 5)*
>
> *"The pattern looks like the readings turned to the right, or laid on their right side." – Student 4 (Question 8, "In a complete sentence, describe what might cause the pattern.")*

Both of these examples describe the pattern in terms of color change. Based on the descriptions provided by the students, it was difficult to determine to what extent their observations reflected a pattern as defined in the task:

> *"A pattern is when something is placed in a way that is not completely random. There is an order to the way things look." – The definition of a pattern included in the assessment.*

The students' descriptions reference colors but the order of the colors is often poorly described. The second example describes the pattern more thoroughly, referring to the key on the side of the map as a reference point about the order. Student 4's use of the word "readings" as a reference to the legend/key on the side of the window is slightly obscure, and would be difficult to interpret without familiarity with the task.

Not all description based errors were related to the description of patterns, students also used imprecise language to describe their claims and justifications. In the following example from question 16, "write a statement answering the scientific question: Where would you predict to find the Southern Bog Lemming in the United States if the average precipitation decreased by 50cm," the student used the correct principles, but inadequate description, to establish the claim.

> *It would travel further south then normal seeing that it likes damp spots.*
>
> *– Student 7 (Question 16)*

The student was indirectly describing information from the question. The question asked where would you expect to find the animal if all locations on the map were getting 50cm (one precipitation range) drier. The students' claim included information from the question to describe where they would expect to find the animal. Though the location was not well described, the student included justification in their claim. This is an example of a very poorly described location where this animal might be found. While the student may have identified the correct location, there was no reference to amounts of precipitation or geographic regions near the expected location.

### *Content Knowledge Errors*

Content knowledge errors were made up of both geography content knowledge and science content knowledge errors. Both science content knowledge and geography content knowledge errors had sub-types of errors. The sub-types of errors with five or more examples are highlighted in figure 15 with blue stripes and a magenta border.

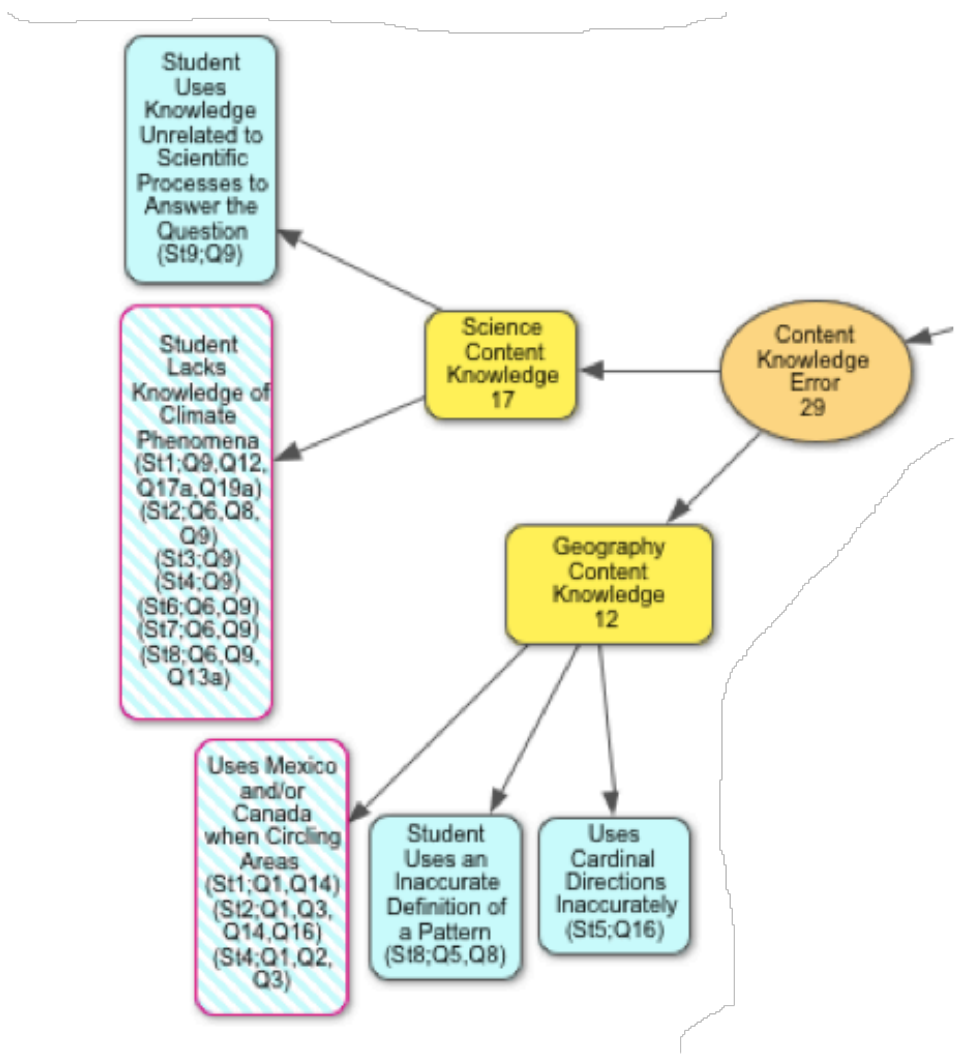



Figure 15. The branch of the complete analysis, figure 2, focused on content knowledge errors. Two types of content knowledge errors were identified, each of which had types of errors, the two types of errors with five or more observations are shown in blue stripes and magenta outline.

Science Content Knowledge

*Student lacked knowledge of climate phenomena necessary to answer the question.*

Some responses represented lack of content knowledge in the area of the task. Students were intentionally not instructed in the content area before the task, and the questions referencing content knowledge were phrased as, "in complete sentences, describe what might cause the pattern" (see Questions 6 and 9, Appendix A). This phrasing allowed students unfamiliar with the cause room for speculation about what might have caused the observed changes. Most responses, showing a lack of science content knowledge, were questions in response to what might cause a pattern. Some examples of these responses are:

> *"I think that Mexico is the reason why it is getting hotter." – Student 1 (Question 12)*
>
> *"I think that different types of time zones/dates cause different temperatures." – Student 2 (Question 6)*
>
> *"because the earth is tilted at its axis the equator is the first thing that gets hit by the sun." – Student 6 (Question 6)*

These responses showed a misunderstanding of the information displayed in the map, with various roots to the confusion. The common trend was lack of understanding about the cause of the variation in annual average temperature and annual average precipitation. Some students' responses reiterated these same points of confusions as justification. The justifications were intended to be less hypothetical, based on the provided information, rather than speculation. Examples of justifications based on scientific misunderstanding were few, but this type of error did occur. Students who used misinformation as a justification also made a similar statement in their description of the cause of a pattern.

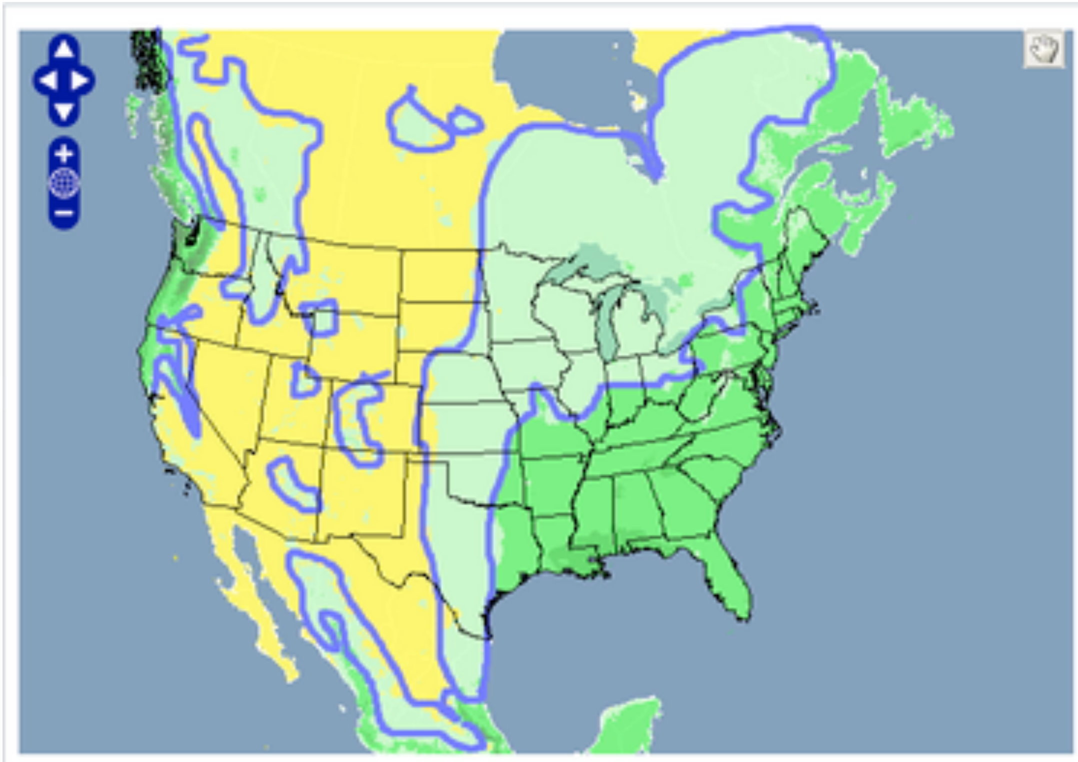
## Geography Content Knowledge

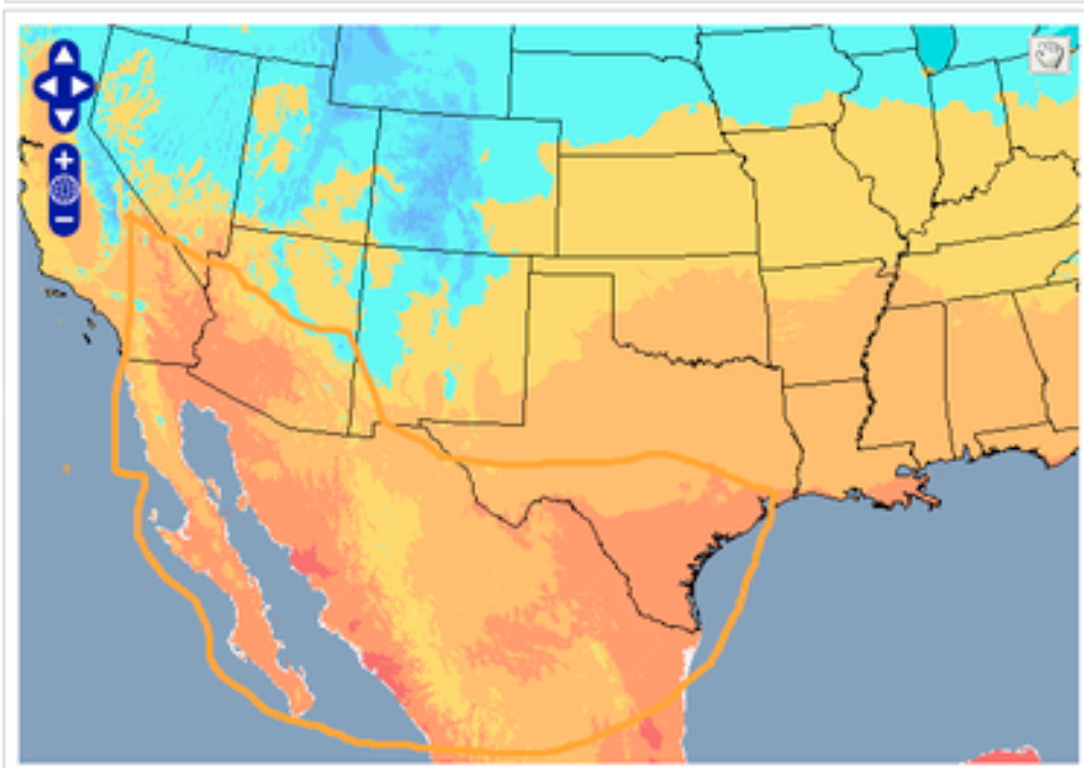
*Student used Mexico and/or Canada when circling areas within the United States.*

Many of the questions stated "in the United States," to limit students' responses to those areas within the lower 48 United States (see Questions 1, 2, 3, 10, 11, 14, 15, 18 in Appendix A). Alaska and Hawaii were not included in the dataset provided to the students. Some students expanded their selection of the correct areas, to include areas outside the United States, as illustrated by the top example, figure 10.

Other responses could have been different had the student heeded the limitations specified in the directions. For example, figure 16 would have had a different response without the inclusion of Mexico. It is unclear what the bounds of the student's response would have been, within the United States, as the circled parts included four different value ranges. Both of these maps, figures 16 and 17, highlight different ways responses can be affected by inattention to specific directions. It is not apparent why the students circled areas in Mexico and Canada and any inquiry would have alerted them to the mistake. One of two problems may have caused this error, either the student forgot to limit the areas to those within the United States, or the student was unaware that the lines showing the border of states also represented the United States border.

This task was designed to draw on a limited amount of content knowledge from students. There was no instruction in task related content knowledge prior to students performing the task. The students' errors provided some valuable insight into their understanding of climate phenomena. Some students' responses revealed confusion about seasons, what causes a place to have a particular average annual temperature and average annual precipitation amount. Of particular confusion to students, was the actual precipitation trends. While this is a complex weather phenomenon, it is interesting to note students' ideas about the origination of precipitation patterns. This is valuable for targeting instruction for understanding of weather and climate. Students' errors from early theories about climate and weather needed correction to help them understand how these concepts are scientifically constructed. A common misconception was the source of seasons and latitudinal differences in temperature. Students often attributed trends in temperature to "elevation". This information makes it clear how confusing a concept such as climate change could be for a student who understands climate as a phenomena related purely to the Earth's distance from the sun, or heat as a function of proximity to Mexico.





Figures 16 & 17. Figure 16 (top), shows a student's response in which the student circled the correct range of values, but included areas within Canada. Figure 17 (bottom), shows a response in which a student included areas within Mexico.

***Science Practices Errors***

Science practice errors were those errors that students made in while participating in the scientific practice of prediction. Though it was expected that some students might make errors while making a claim, none were observed. The six sub-types of prediction errors were all related to students' use of justification to support their predictions. Three of the sub-types of errors had five or more examples and are described in this section. These subtypes are shown below (figure 18) and highlighted with blue stripes and a magenta border. Each of these types of error is elaborated in the subsequent sections.

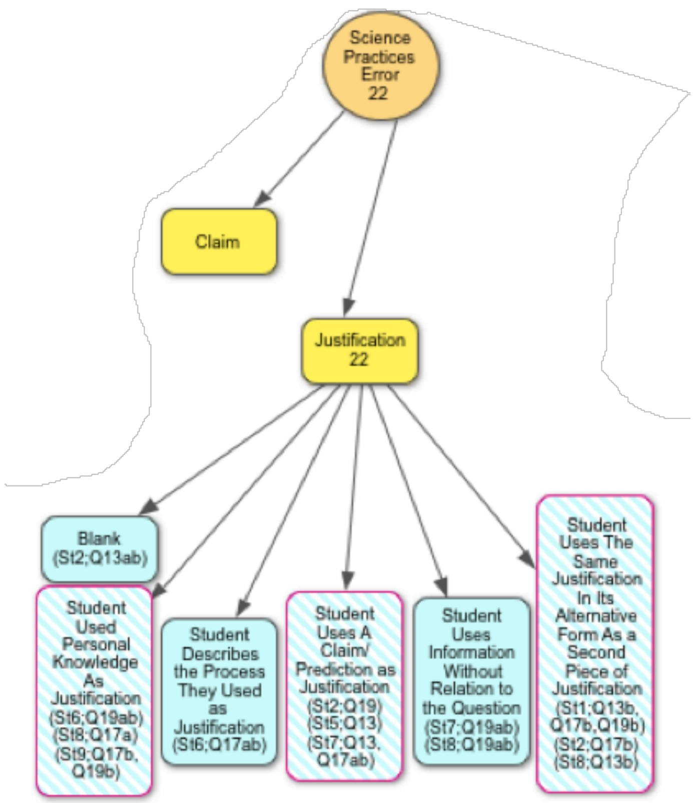



Figure 18. The branch of the complete analysis, figure 2, focused on science practices errors. Two categories of science practice errors were identified only one of which had observed errors, justification errors, the types of errors that were observed are shown in blue the types with five or more observations are shown in blue stripes and magenta outline.

## Justification Based Errors

Students made a variety of errors in supporting their predictions with facts or scientific principles. Since no scientific principles were presented in the task, they needed to use either the facts presented, or other scientific principles they were familiar with to support their predictive claims. This resulted in a variety of errors, six types in total, four of which had five or more examples to illustrate the error.

### *Student described the process they used as justification.*

A similar problem is present when interpreting some justification pairs provided by students. Students often referred to the map as justification without providing reference to a particular aspect of the map that they were looking at:

*"1) Looking at the map and comparing answers from given data*

*2) Putting correct colors with the maps." – Student 6*

While the student describes the process used to complete the task, they do not provide any of the details that would support their response. The lack of detail in the description makes it impossible to tell which colors the student was looking for, or which data they used to compare with the map. In the evidence regarding the location of a species' habitat, neither the species, nor the species' habitat preferences are referenced in the evidence. The justifications presented are not wrong, but they lack the precise detail necessary to tie the evidence from the map to the claim made about a species' habitat. This example shows the middle ground between a justification error and an error caused by the level of detail in the description. Justifications are, in a sense, descriptions of the process the student used to answer the question, but the key component of the justification is the reference to the facts supporting the claim. The process used to answer the question, absent the justification used to answer the question, does not back a claim.

*Student used personal knowledge as a justification.*

Some students used unrelated personal knowledge as justification for their predictions. A piece of unrelated personal knowledge is a piece of information that was not discussed in the presented information. Some examples include:

*"The areas where tornados are most common." – Student 9 (Question 17)*

Neither tornados nor more general natural disasters were discussed in the task. Weather and climate phenomena were presented in terms of annual average temperature, and the potential for change in those averages was presented. It is unclear how this student related this information to tornados. The student did seem to be using the locations where tornados occur as a point of reference, as though the location of the extent of high tornado occurrence was common knowledge. This was the complete statement written as a justification. Another example demonstrates that if a student referenced information insufficiently, it became unclear if the student was basing their judgment on personal knowledge or information provided in the task. In the following example, the student uses vague terms, making it unclear whether the information used is from the task.

> *"For the [justification][1] the bog lemming likes the heat so with less rain level and in dry spots it would like to stay there." – Student 8 (Question 17)*

The student refers to the animal as "likes the heat" and "dry spots." These phrases are referential to a standard, which is not clearly defined. If there wasn't a clear description of cold or cool, then likes heat refers to an undefined range. While this may appear to focus on minute details the students were given a range of values and colors on a map. This student translated those ranges into a personal value system, not clearly specified, and used that system as a justification for the prediction.

*Student used a claim or prediction as a justification.*

Some students used a claim as justification for a different claim. For example, a student might say, I would expect to find the animal in the northwest United States, and support that claim with further explanation of what might happen to the animal. An example from a students' work is:

> *The plants and animals would die and travel further like the bog lemming to find food. – Student 7 (Question 17)*

This student further speculated about what might happen to other animals, after having made a claim about what they expected to happen to the bog lemming.

[1] Justification is used in this paper. Students were asked to support their statements with "backings". A different word was presented to students with the same definition. This slight change in language is used for consistency and clarity.

Other students used an alternative description of the same claim as the justification for that claim. This student used the map as a justification for the information circled on the map.

*The other [justification] would be the map that shows the new areas of the Southern Bog Lemming. – Student 5 (Question 13)*

This student seems to be citing the map as evidence for the claim, but without specific information from the map to support the claim, it is a restatement of the claim as a justification.

*Student used the same evidence in its alternative form as a second piece of evidence.*

Some students used two pieces of evidence, stating the same thing in reverse. The two pieces of evidence referred to the same fact. The second piece of evidence provided no additional information.

*"Southern states are always hotter because they are on a lower elevation*
*Northern states is always cold because they are on a higher elevation"*
*– Student 1 (Question 19)*

The second statement is a restatement of the first piece of evidence from an alternative perspective. That change in perspective is characteristic of this type of error. Students generally are not restating the identical piece of information. The change of perspective in their approach to the same information may mean that they do not grasp that it is the same concept.

Justifications frequently took the form of a description of the process used, a claim, or personal knowledge. This could represent the students' lack of content knowledge appropriate to these contexts, although the necessary supporting content knowledge was provided. Alternatively, it could be that students do not adequately understand the type of information needed to support their claims. This would seem to be a better fit for the evidence, from the research, since the justifications related errors were due to the inclusion of inappropriate types of justifications. A counter-point would be that some justifications errors were classified as description errors. There were only two errors (one each by two students) in which the student inadequately described an appropriate justification. Seven of the nine participants made a justification error of some sort, pointing to the broad need for instruction. Justifications errors underline the need for targeted instruction and scaffolding to help students understand the appropriate support for their predictions.

## Discussion

Each of the four major types of error identified in this study offer potential implications for instruction and curriculum development. The following discussion focuses on three of the four types of error. It is assumed that the ten observations related to directions errors could be eliminated with increased familiarity with the format of the practices presented, and careful rewording of the assessment questions causing difficulty. Focusing on science practice errors and data practice errors in the context of the role of content knowledge allows consideration of how instruction in these practices relates to the integration of content and practices.

## Scientific Practice Errors

Mistakes made in the construction of the predictions, seemed to fall into two categories: uncertainty about what qualified as a justification, or reliance on information from somewhere other than the assessment. The six types of error can be categorized into one or both of these two areas. Previous research has found that students have difficulty selecting evidence for explanations when they are less familiar with the content area (McNeill & Krajcik, 2007; Sandoval, 2003; Sadler, 2004). In the McNeill and Krajcik (2007) study, in particular, this seemed attributable to a less than confident scientific understanding of the statement the student uses as reasoning. Similarly, Sadler's (2004) review of socio-scientific explanation literature found that most studies identified that participants were able to support their claims better in areas where they possessed more content knowledge. Students in this study had the most difficulty with supporting claims in their predictions, which do not require reasoning.

In Sandoval's (2003) study of a technology based framework for constructing explanations, students were able to add physical connections between their descriptions of evidence and the data they used to make that statement. Students used graphs and histograms as data from which they selected elements to serve as evidence. The interface allowed students to cite the information (from the graph or histogram) they used to draw conclusions. Students in this study had a limited amount of information from which to draw, since instruction or a curriculum associated with the material was not provided. Students in the study seemed to rely more on prior knowledge to support their claims than to use information given in the task itself. This finding is consistent with the findings of Schunn and Anderson (1999). Both studies show that novices who are uncertain of the criteria that justifications must meet to support a scientific claim tend to use personal experience or theories in support of their claims.

In students' development of justified predictions, there is no element of reasoning to ensure that the evidence is tied to the claims. This changes the approach from those considered in the literature previously presented. Reasoning is generally considered to be a scientific principle or definition that supports the claim the student is making. It is sometimes referred to as a fact that links the evidence to the claim. What remains unclear is whether the lack of reasoning in this task made the selection of justification more difficult for students. For some, the reasoning statement could serve as a filter through which data is excluded from consideration as evidence. This exclusion can occur too early in the process, resulting in the exclusion of disconfirming evidence (Kuhn, 1993). In terms of developing justified predictions it may be that the lack of reasoning, limits the students' ability to exclude evidence based on fit with the claim and reasoning.

## Data Practices Errors

There are many circumstances that could cause the difficulty that students had with data practices errors. Teachers are often unprepared or unable to help students with the selection of evidence and the examination of data. Students in schools that use mainly traditional teaching methods might not be using representations like maps, in order to create scientific explanations and predictions (Vekiri, 2002). Static maps are often the graphics presented in texts and support the students' understanding of knowledge being presented (Vekiri, 2002). Students may have been expecting the maps to tell them information, rather than constructing knowledge from the map. Though their experience with inquiry-based curriculum may have prepared them to interpret and analyze the maps, their responses did not indicate an advanced understanding of the practices. Schools' infrequent use of data based maps would lead to an underdeveloped data practices emphasized in this task.

Some might argue that because students in this assessment were turning on and off data layers in order to answer different questions, they were constructing the maps and the knowledge related to the data presented. This would be an inaccurate characterization of the tasks represented in this study. Participants in this study turned on and off data layers in correspondence with the directions presented to them. They were required to respond to questions asking for their interpretations of the static representation they had created, or answer questions intended to draw specific information from the representation. In this way, these tasks were more closely linked to the school use of representations, though the major difference was in the purpose. These representations were designed for participants to use for the purpose of making claims, but not for communicating claims to the participants.

Due to the design of the data practices tasks, content knowledge deficiency seems to be largely independent of the errors classified as data practices errors. While content knowledge played an important role in students' construction of explanations, the data practices focused tasks required very little background knowledge, focusing on interpretation of the presented data. Future assessments should give students a particular focus for their analysis. In these tasks students were responsible for responding to the associated questions, but few students had a sense of the broader question being answered by the analysis. In the data practices errors, students made mistakes in their interpretation of the representation absent a specific scientific context. Studies (i.e. Roth, 2002) have shown the differences between an expert and novice's interpretation of context dependent representations. While a novice might be able to interpret the basic changes in a representation, the expert interprets the representation as a real world story (Roth, 2002). Students looking at the representations of climate used in this assessment might interpret the visual in the most basic sense, but even without the context of content knowledge, basic interpretations should still be possible, assuming that students have the underlying process skills. Students' responses on this assessment show that many students lack the underlying process skills necessary for a basic interpretation of the representations.

This is not to say that content knowledge is unimportant in data practices. Roth and others' work (Roth, 2002; Shah et al, 1999) show that experts unknowingly draw on a significant amount of content knowledge when interpreting representations. While content knowledge was raised as an important determinant of students' abilities for both the construction of explanations (Sadler, 2004; Sandoval, 2003; McNeill & Krajcik, 2007), and for the interpretation and analysis of representations (Vekiri, 2002), the initial interpretations in this study were within the ability of the participants. These identification tasks relied on general scientific knowledge and on the knowledge of a scientific practice. Since students had difficulty with these practices in the absence of new content knowledge, this indicates a fundamental lack of experience with the scientific practice.

Participants in this study sometimes struggled with the basic interpretations that did not require content knowledge. When they moved from the recognition of basic features to interpretations with descriptions, their lack of related content knowledge became more apparent; these errors were classified as content knowledge related errors.

Students need to be able to interpret representations accurately without the context of underlying content knowledge. It is assumed though, that interpretation without content knowledge will result in a response without the same level of knowledge and detail.

### Content Knowledge

In a different practice, drawing inferences from data, students felt comfortable making inferences about the cause of patterns based on the knowledge they had of other processes, for example in the question about the cause of the pattern of precipitation variation. This same process, using knowledge to support claims, is what was expected for the predictions students made. Many students had difficulty using these same practices in a novel situation, a prediction with claim and justifications.

Content knowledge is an important part of progressing with the use of each of these types of practices, both in the construction of justified predictions and in data practices. These results show that students have difficulty interpreting geographically variable data even in the absence of complex content knowledge. The data presented demonstrates that students struggle with the fundamental elements of some of these data and scientific practices. This may become more taxing for students when novel content knowledge is presented in conjunction with new practices.

## Conclusions

This task was not intended to characterize students' capabilities or capacities for learning. The focus of this task was on characterizing students' unassisted capabilities in performing an interpretation and analysis task with geographically variable data. Through the identification of unassisted abilities, scaffolds can be developed to help students as they learn the practices and content emphasized in the curriculum. While this analysis focused on errors that students made, it is important to note that some students did complete the task with relatively few errors. Many students who completed the assessment would benefit from scaffolding and instruction to better understand data and scientific practices.

This research identified four areas where students can get stuck using processes that will not lead them to the correct answer. These categories of errors, task directions, data practices, science practices, and content knowledge, describe the broad areas of difficulty that students had completing an increasingly more difficult assessment task. While it can be assumed that a traditional classroom task would have more instruction and context provided with it, this task required little prior knowledge as a trade-off to not having associated instruction. The difficulties that students had with this task provide valuable insight into the areas where instruction should be focused when incorporating a similar task into a curriculum.

Instruction seems to be particularly necessary for helping students understand what the data they are working with represents and also how to support their scientific claims with evidence. The difficulty that students had with basic data interpretation tasks leads to the conclusion that students need instruction for developing scientific practices in concert with content knowledge. Students struggled with the interpretation and analysis tasks in part because they lacked the content knowledge to focus their observation. It is rare that an individual is asked to analyze a dataset without a scientific question to answer. As students advance through the grades, if they continue to struggle with scientific practices, adding more complex content knowledge might not help them to develop their abilities to use scientific practices. More research is needed to look at the kinds of classroom activities and instruction that will assist students in developing their data practices. Future research could focus on the question: what types of scaffolds support students' use of geographically variable data as evidence in support of a scientific prediction? In future studies it would be necessary to work with students approaching a specific classroom task, to better understand how students with a specific scientific question approach the interpretation and analysis of data, as well as the selection of evidence for their predictions. Through this research question the types of scaffolds that support students' data and scientific practices could be identified and improved through work on an in class investigation.

# Bibliography

American Association for the Advancement of Science (AAAS). (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Bausmith, J. M., & Leinhardt, G. (1998). Middle-School Students' Map Construction: Understanding Complex Spatial Displays. *Journal of Geography, 97*(3), 93-107.

Bednarz, S. W., Bettis, N. C., Boehm, R. G., deSouza, A. R., Downs, R. M., Marran, J. F., et al. (1994). *Geography for Life: National Geography Standards 1994*. Washington, D.C.: National Geographic Research & Exploration.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: a practical guide. *Journal of the Learning Sciences, 6*(3), 271-315.

College Board. (2009). *Science: College Board Standards for College Success*. Princeton, NJ: College Board.

Davis, E. A., & Linn, M. C. (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. *International Journal of Science Education*, *22*(8), 819-837.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academy Press.

Edelson, D. C., Gordin, D., & Pea, R. D. (1997). *Creating Science Learning Tools from Experts' Investigation Tools: A Design Framework*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: MIT Press.

Gordin, D. N., Polman, J. L., & Pea, R. D. (1994). The Climate Visualizer: Sense-Making Through Scientific Visualization. *Journal of Science Education and Technology, 3*(4), 203-224.

Inspiration Software, Inc. (2011). Inspiration Software (Version 9) [Software]. Available from http://www.inspiration.com/

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education*, *77*(3), 319-337

Lee, H. S., & Songer, N. (2003). Making authentic science accessible to students. *International Journal of Science Education*, *25*(8), 923-948.

Leinhardt, G., Stainton, C., & Beusmith, J. M. (1998). Constructing Maps Collaboratively. *Journal of Geography, 97*, 19-30.

Linn, M. C. (2006). The knowledge integration perspective on learning and instruction. In R. K. Sawyer (Ed.), *The cambridge handbook of the learning sciences* (pp. 243-64). Cambridge University Press New York.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting Students' Construction of Scientific Explanations by Fading Scaffolds in Instructional Materials. *Journal of the Learning Sciences, 15*(2), 153-191.

McNeill, K. L., & Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In M. C. Lovett & P. Shah (Eds.), *Thinking with data.* (pp. 233-65). New York, NY: Lawrence Erlbaum Associates

Miles, M. B., & Huberman, A. M. (1994). *An Expanded Sourcebook: Qualitative Data Analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.

National Research Council. (2011). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council (NRC). (1996). *National Science Education Standards*. Washington, DC: National Research Council.

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3 ed.). Thousand Oaks: Sage Publications.

Roth, W. M. (2002). Reading graphs: Contributions to an integrative concept of literacy. *Journal of Curriculum Studies*, *34*(1), 1-24

Ruiz-Primo, M. A., Li, M., Tsai, S.-P., & Scheider, J. (2010). Testing One Premise of Scientific Inquiry in Science Classrooms: Examining Students' Scientific Explanations and Student Learning. *Journal of Research in Science Teaching, 47*(5), 583-608.

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching*, *41*(5), 513-536. doi:10.1002/tea.2000

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *Journal of the Learning Sciences*, 5-51

Schunn, C. D., & Anderson, J. R. (1999). The generality/specificity of expertise in scientific reasoning. *Cognitive Science*, *23*(3), 337-370

Shah, P., Mayer, R. E., & Hegarty, M. (1999). Graphs as Aids to Knowledge Construction: Signaling Techniques for Guiding the Process of Graph Comprehension. *Journal of Educational Psychology, 91*(4), 690-702.

Songer, N. B., Kelcey, B., & Gotwals, A. W. (2009). How and When Does Complex Reasoning Occur? Empirically Driven Development of a Learning Progression Focused on Complex Reasoning about Biodiveristy. *Journal of Research in Science Teaching, 46*(6), 610-631.

Songer, N. B., Shah, A. M., & Fick, S. (In Press). Characterizing teachers' verbal scaffolds to guide elementary students' creation of scientific explanations. *School Science and Mathematics.*

Telestream, Inc. (2011). ScreenFlow Software (Version 3.9) [Software]. Available from http://www.telestream.net/screen-flow/overview.htm

Toulmin, S. (1958). *The Use of Argument*. Cambridge, England: Cambridge University Press.

Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology Review*, *14*(3), 261-312

Vygotsky, L. S. (1978). *Mind In Society*. Cambridge, MA: Harvard University Press.

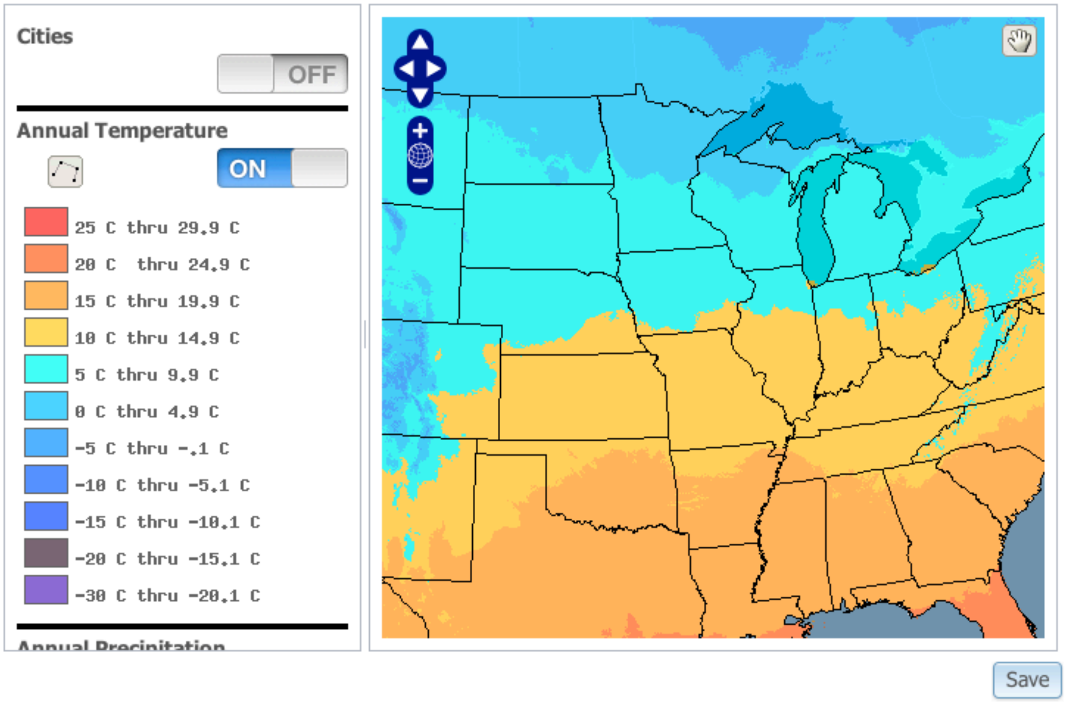
White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, *16*(1), 3-118

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the Scientific Model: Model-Based Inquiry as a New Paradigm of Preference for School Science Investigations. *Science Education, 92*(5), 941-967.

# Appendix A. The Assessment Task

## Example Item

Michigan Map



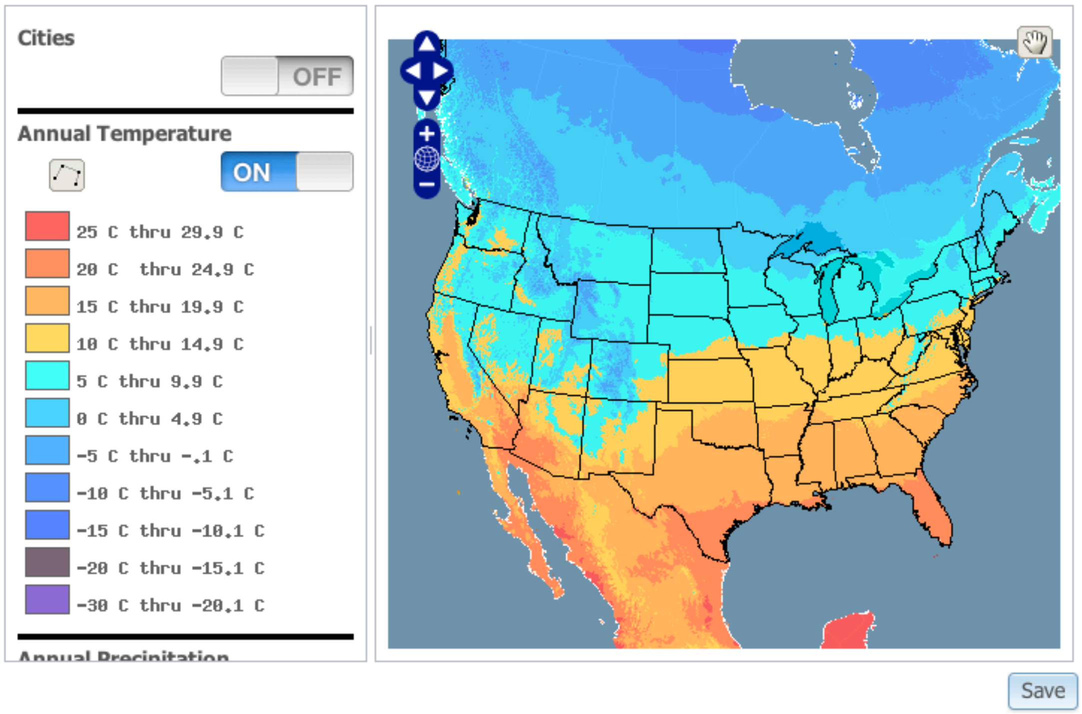

To Draw a Border:
Either the Temperature or Precipitation Layer needs to be turned "ON". Then you need to click on the box that looks like its a game of connect the dots. Now, when you click and drag on the map a line will appear. To draw a border, click and hold the mouse button, dragging the pointer around the edge of the state of Michigan.

**Please draw a border around the state of Michigan on the above map.**

When you are finished, click "Save" under the map, and then "Next" in the top right corner.

## Map 1

Map: Hottest Temp ID

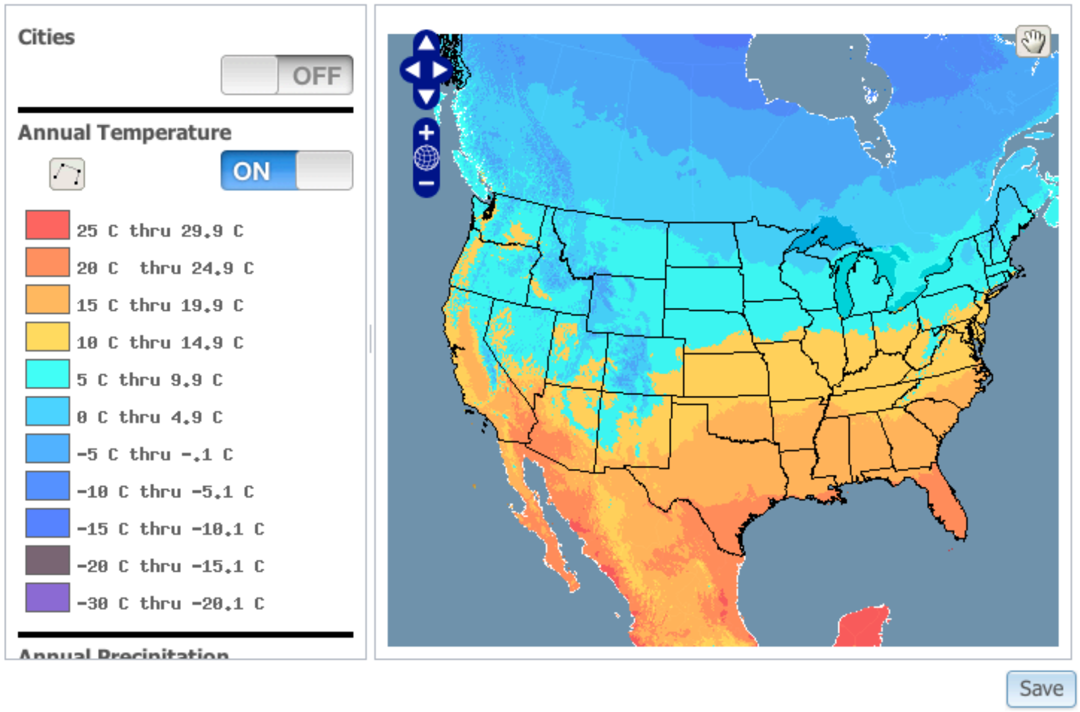



Using the map above with the temperature layer turned on, answer question 1.

1) Draw a border around the areas within the United States with the hottest average temperatures.

## Map 2

Map: Temp Location

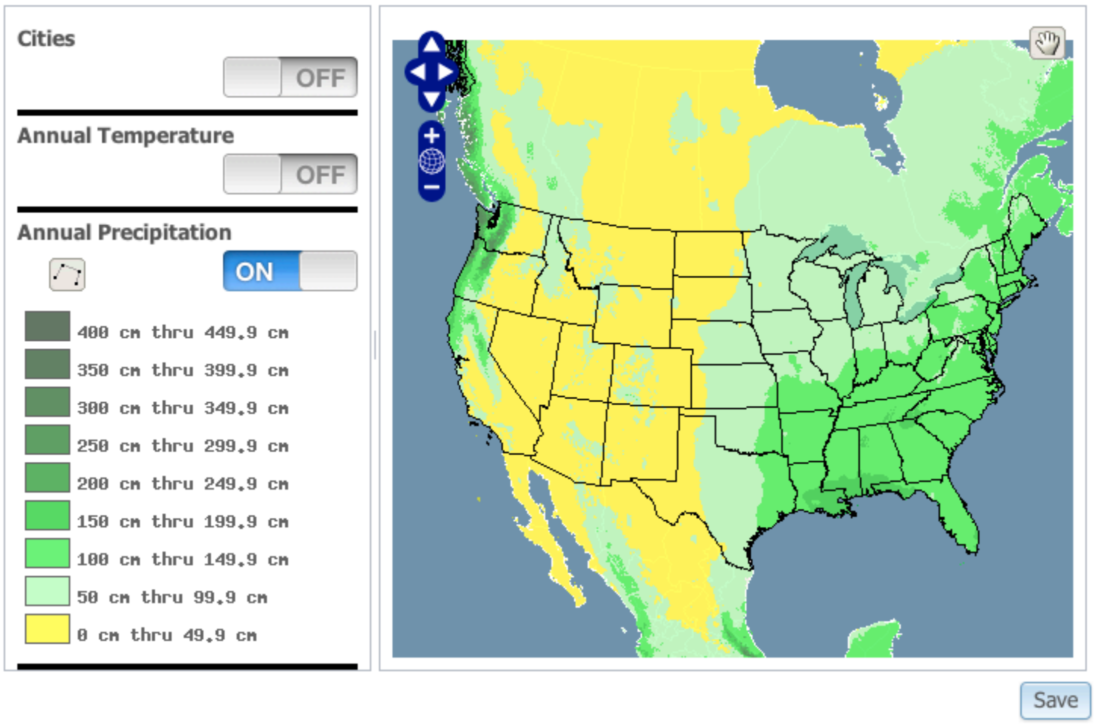



Using the map above with the temperature layer turned on, answer question 2.

2) On the map, draw a border around all areas within the United States that have average temperatures between 15°C and 19.9°C.

## Map 3

Map: Precip Location

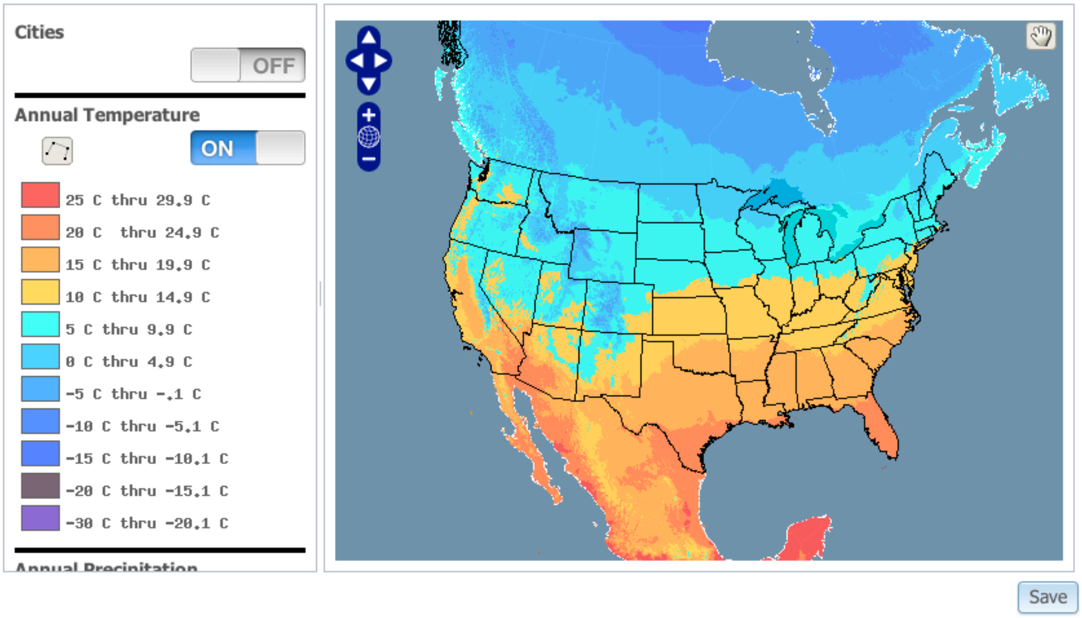



Using the map above with the precipitation layer turned on, answer question 1.

3) On the precipitation map, draw a border around all areas within the United States that have precipitation between 50cm and 99.9cm of annual precipitation.

## Map 4

Map: Temp Pattern ID




Use the map with the temperature layer turned on and the explanation of a pattern to answer questions 4-6.

A pattern is when something is placed in a way that is not completely random. There is an order to the way things look.

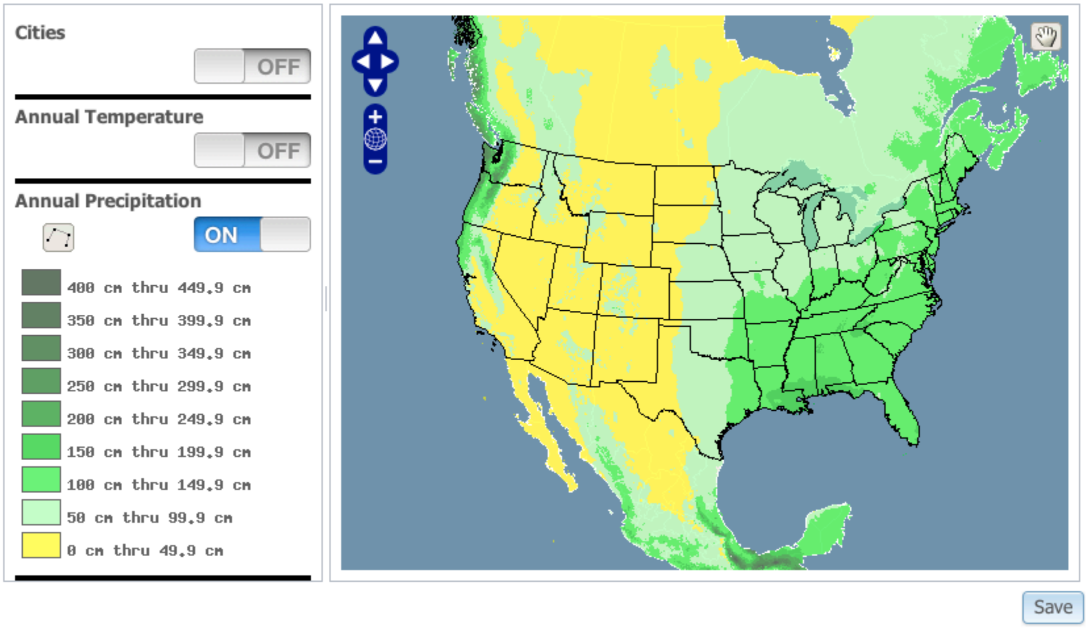
4) Draw a border around an area with a pattern.

5) In a complete sentence, describe how the pattern looks.

6) In complete sentences, describe what might cause that pattern?

## Map 5

Map: Precip Pattern ID




Use the map with the precipitation layer turned on and the explanation of a pattern to answer questions 7-9.

A pattern is when something is placed in a way that is not completely random. There is an order to the way things look.

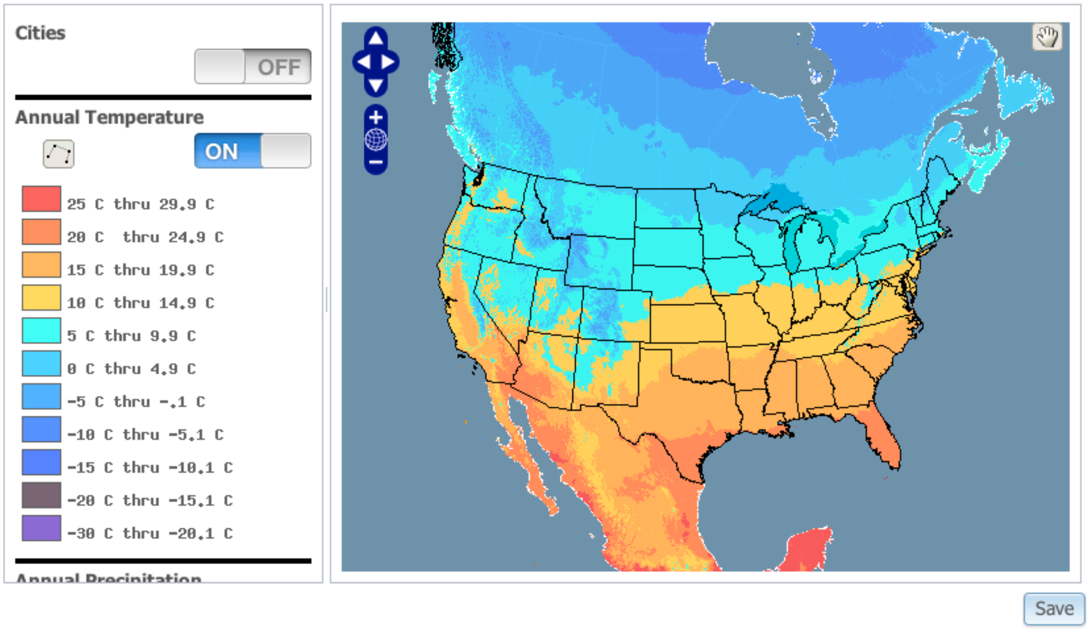
7) Draw a border around an area with a pattern.

8) In a complete sentence, describe how the pattern looks.

9) In complete sentences, describe what might cause that pattern?

## Map 6

Bog Temp Prediction




Using the map above with the temperature layer turned on, answer questions 10-13.
The Southern Bog Lemming likes temperatures between 10°C and 19.9°C.

10) Draw a border around the areas where you would expect to find the Bog Lemming in the United States, on the map above.

11) Draw a border around the areas where you expect to find the Southern Bog Lemming if the United States were getting hotter, 5°C warmer everywhere.

12) Write a statement answering the scientific question: Where you would predict to find the Southern Bog Lemming in the United States if the average temperature increased by 5°C?

Backing: data or a scientific concept or definition that supports your answer to the scientific question

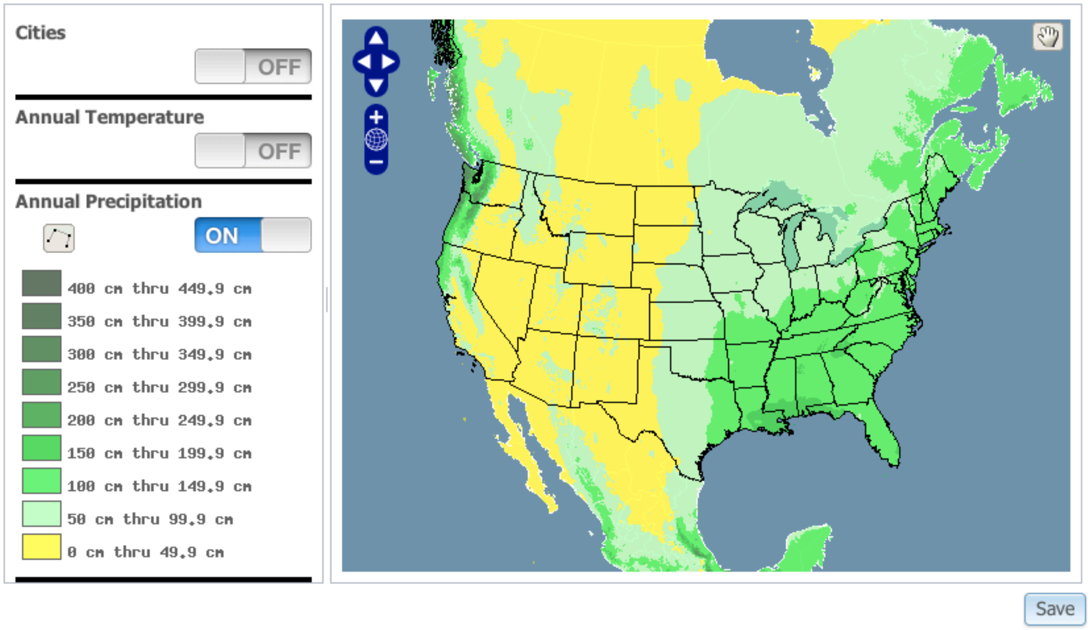
13) What two backings do you have to support your prediction? Describe them in complete sentences.

1)

2)

## Map 7

Bog Precip Prediction




Using the map above with the precipitation layer turned on, answer questions 14-17.
The Southern Bog Lemming likes precipitation between 50 and 149.9 centimeters.

14) Circle the area where you would expect to find the Bog Lemming in the United States, on the map above.

15) Circle the areas where you expect to find the Southern Bog Lemming if the United States were getting drier, 50 centimeters less precipitation everywhere.

16) Write a statement answering the scientific question: Where would you predict to find the Southern Bog Lemming in the United States if the average precipitation decreased by 50cm?

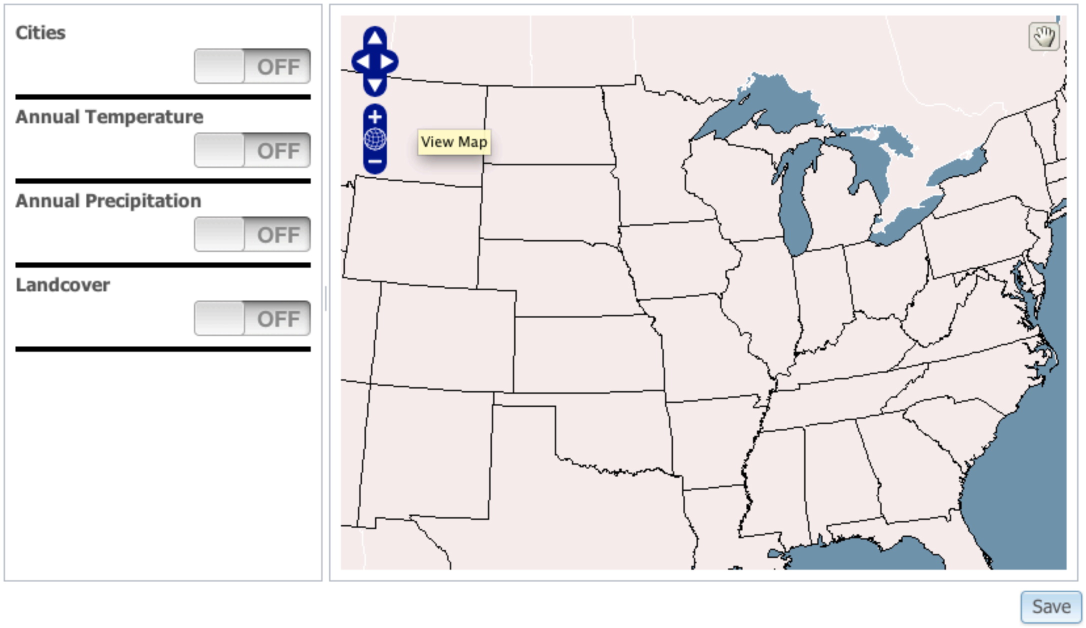
Backing: data or a scientific concept or definition that supports your answer to the scientific question

17) What two backings do you have to support your prediction? Describe them in complete sentences.

1)

2)

## Map 8




Turn on both the temperature AND precipitation layers to answer question 18.
Use the borders you created on the last two maps, which are visible when you turn on the temperature and precipitation layers to answer the following questions.

18) Write a statement answering the scientific question: Where would you expect to find the Southern Bog Lemming in the United States under the present temperature and precipitation conditions?

---

Backing: data or a scientific concept or definition that supports your answer to the scientific question

19) What two backings do you have to support your prediction? Describe them in complete sentences.

1)

2)