Approved MARK HARROWER

Advisor Title ASSOC. PROFESSOR

Dept. of Geography

Date 05/10/10

On Varying the Detectability of Symbols to Encode Data Uncertainty

by

Daniel P. Huffman

A thesis, submitted in partial fulfillment of
the requirements for the degree of

Master of Science
Cartography and Geographic Information Systems

at the

University of Wisconsin – Madison

2010

In Gratitude

This thesis represents the pinnacle of my academic career; I have no designs on pursuing higher studies, and plan to recede from these hallowed, labyrinthine halls into a life outside the land of postnominal letters. The Academy will not, I suspect, long recall me once absent, and so these words are in many ways my last testament. I hope you will indulge me, gentle reader, as I shout a few words from the summit, small as it may be.

Two years ago, I was adrift, mired in a deep depression, directionless in life, relationships in tatters, and struggling in a graduate program with which I was incompatible. And then some kindly people in Geography took a chance on me, a man who had floundered in his first attempt at a Master's a few blocks away and whose personal essay said little more than "I really like maps." I do not know what would have happened without this opportunity, and the journey that this document represents. I would be in a very different and, I fear, very dark place. In this building I have found a new passion which drives me, and a welcoming community to share in the joy.

Through much of these two years, I have been under the tutelage of Mark Harrower, whose guidance, wit, and knowledge have been invaluable. First impressions are critical, and I leapt into cartography very much on account of the skill and passion with which he taught me my introductory course. Mark also hired me into the UW Cartography Lab, giving me a home away from home – a community of friends and colleagues who have taught me much of my art and kept me laughing. To Kevin, Tim, Rich, Colleene, Dan, Ben, and of course, our esteemed Associate Director Tanya: I am not sure that I could have made this journey without all of you around me. I am much the richer for our time together.

I have accumulated many other debts in my academic career, too numerous to give a fair accounting. I will leave you with a scattered few of the most prominent ones, and I hope those whom I have left off will be forgiving.

To Jim Burt, who taught me statistics twice – once in class, and once again when I'd forgotten the class and needed to write this thesis.

To the generous people of my far-off home of Kalamazoo. I hope that, in some small measure, I have made good on the many opportunities you gave me to make it this far. I wanted to make you all proud.

And to my friend Greg, who told me that maybe if I liked maps so much, I should think about getting a degree or something in how to make them.

Daniel P. Huffman
Madison, WI
5th May, 2010

Introduction

The data which make their way on to a map are never perfect. There is an unavoidable imprecision in measuring the elevation of a mountain, for example. When mapping census data there is an error arising from the assumption that the sample population represents the whole. Many types of these problems can be found underlying the data on to a map, and they go by many different terms – such as error, imprecision, unreliability – but the word "uncertainty," as MacEachren (1992) suggests, is a good blanket term to cover the many different ways in which we can never be sure of the degree to which our maps reflect reality

While cartographers can and do make maps which display the uncertainty of their data, there has until recently been little systematic investigation of how to symbolize it, in order to determine which methods convey this information to a reader most effectively. The bulk of previous cartographic research, such as MacEachren (1992), MacEachren *et al.* (1998), and Leitner and Buttenfield (2000), has focused on how to apply specific visual variables to uncertainty representation. Visual variables are properties of an image, such as the size or shape of a mark, that can be manipulated systematically to convey information. However, there is another, more fundamental property of images which has not been yet been examined as a potential carrier of uncertainty information in maps: *detectability*. The detectability of a symbol is simply how easily it can be noticed by a viewer, and it depends on its size and the brightness contrast it has with its background (Riggs, 1965). Both size and brightness (often referred to as color value) are visual variables, and while experimentally manipulating these to represent uncertainty, previous researchers have at the same time been altering detectability, though they have not explicitly addressed it. Thus, it is not clear if the experimental results of earlier authors can be attributed simply to the fact that they changed the size or color value of symbols, or if in fact their results are better explained by how those symbols differed from their visual context. Extracting useful guidance from prior experimental work requires an answer to this question. This thesis examines the role of detectability – of visual context – in creating a perception of uncertainty in map readers.

In order to examine potential relationships between the *physical* phenomenon of detectability (which relies on changes in light striking the retina) and the *cognitive* notion of certainty, this thesis puts forth two hypotheses:

1) In absence of a legend to guide them, map readers will tend to interpret more detectable symbols as representing more certain data, rather than less certain data.

2) Readers will more accurately extract uncertainty information from maps that encode more certain data by more detectable symbols than they will from maps that encode certainty with the opposite scheme (less certain is more detectable).

Background and Literature Review

A brief note on the term "uncertainty" is warranted. Uncertainty itself is a rather broad term, which can encompass a variety of types of data quality problems. For example, a data set may be out of date, or from an untrusted source, or have known measurement inaccuracies. MacEachren *et al.* (2005) provided a synopsis of several typologies of uncertainty. I will leave aside the details of what is, in truth, a complex topic with its own literature. Roth (2007) has found that the type of uncertainty which was displayed on a test map did not influence the decisions made by users based upon that map, though it did affect the speed and confidence of their decisions. Therefore, the particular details of how uncertainty is generated, propagated, conceptualized, and described should have no particular bearing on this research, which is instead how it is represented.

Visual Variables

Jacques Bertin (1983) proposed that static visual communication could be broken down into eight "components of the graphic sign-system," which he referred to as visual variables: the two dimensions of the plane, size, value, texture, color, orientation, and shape. By color, Bertin meant what we might more commonly refer to as "hue" – a color's dominant spectral component (blue, red, green, violet, etc.). Value Bertin describes as "the various degrees between white and black," which he treated as a combination of the brightness of a color and its saturation. Value will be referred to as "color value" throughout, to avoid confusion of the term with the concept of numerical values. Through variations in these eight visual properties, a visible mark encodes information. Cartographers employ the dimensions of the plane – a mark's position on the page – to convey spatial locations, and so rely on the other six variables to communicate thematic information. Bertin called these other six the "retinal variables."

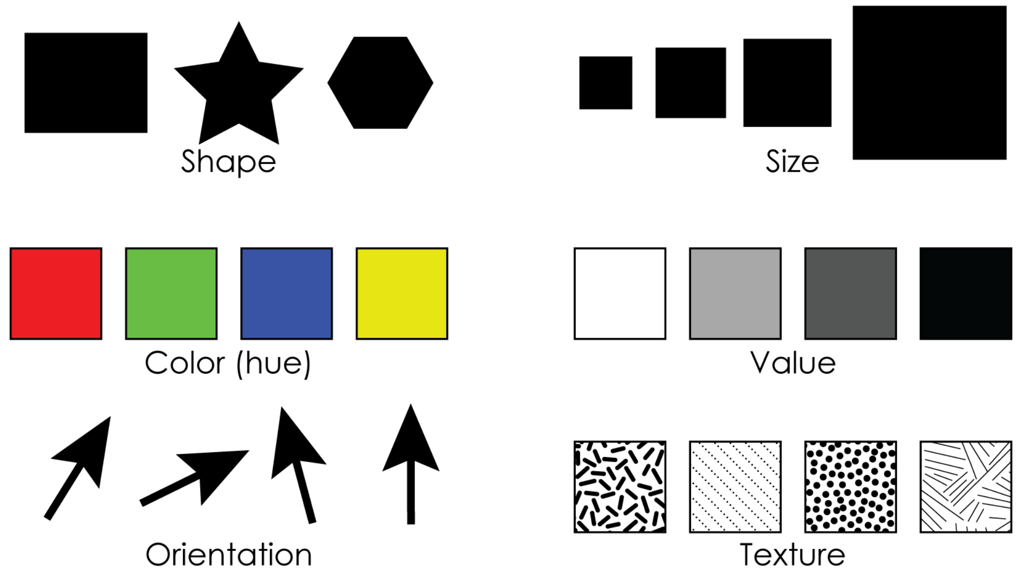

FIGURE 1 – Bertin's six "retinal variables."

Many of the works covered in the following pages have taken the idea of visual variables as a starting point – asking which specific ones would be most suited to conveying uncertainty information. Edwards and Nelson (2001), for example, tested maps in which more certain data were shown by darker colors on a white background map. However, there was no test of the effect on readers if the *less* certain data were shown by darker colors. A common thread running through the works discussed below is that authors assume that more certain data should be made to stand out more on the map, though it is generally only an implicit assumption. Its usefulness as a guide to symbol choice has yet to be tested, and the aim of this thesis is to experimentally test whether visual context is a relevant factor in reading uncertainty off of maps.

Early Technique Proposals

Most of the literature discussion on uncertainty representation has taken place in the past two decades, though the concept of encoding data quality was not new. In 1942, Wright discussed techniques which were already in use for representing uncertainty in the mid 20th century: adding notes in the margin of maps, labeling specific features with data quality warnings, adding inset maps which assess reliability of the main map (reliability diagrams), and using broken contour lines for uncertain values. However, the subject was otherwise absent from the literature until the early 1990s, when a series of technique proposals came out. These were strictly theoretical, rather than based on user-testing, which would follow primarily in the several years after.

Alan MacEachren (1992) suggested that matching uncertainty to Bertin's (1983) six retinal variables was an "important representation issue for visualization of uncertainty." All of them, he said, could be used reasonably to convey uncertainty – size and value for numeric assessments of uncertainty; hue, shape, and orientation for nominal (non-orderable) uncertainty information. The presence or absence of a texture might be most useful in a binary classification of "certain enough" or "not certain enough." In addition to Bertin's six, MacEachren also brought up two non-Bertin variables: saturation and focus, as possibly useful. The former, first proposed as a visual variable by Morrison (1974), is "arguably the most logical" variable for indicating uncertainty, though MacEachren does not justify this claim. He proposes employing it such that the most certain data has a pure hue, and the least certain is washed out and grey. He does not state whether or not it would be necessary to change this scheme based upon the background color of the symbol whose saturation is being varied, or whether he believes this rule can be applied absent of visual context.

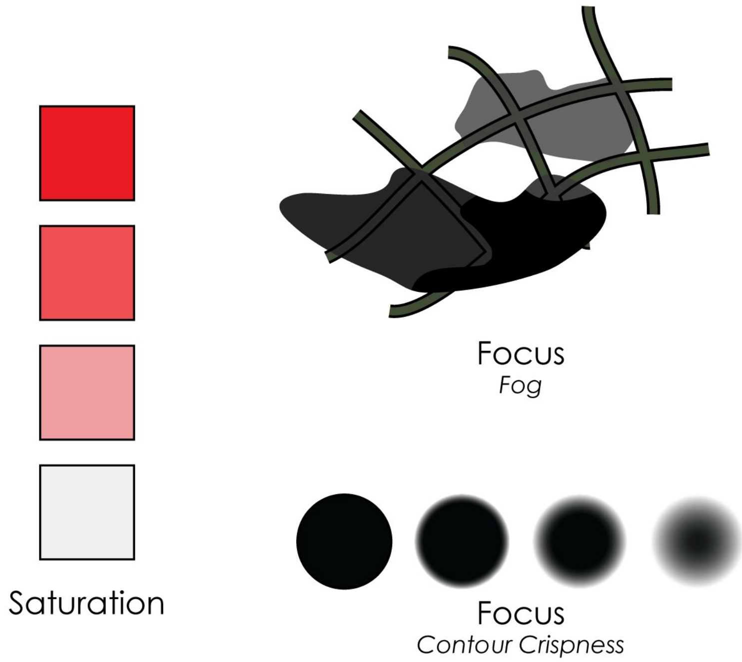

Figure 2 – Additional visual variables: saturation and focus.

Focus, which is a broad variable MacEachren credits David Woodward for inventing but not publishing, involves making information appear less clearly. He suggests several different instantiations of focus. Marks based on uncertain information could be given blurred lines or fuzzy fill textures, or maps could be covered by a "fog" of varying transparency, allowing more certain data to show through and obscuring the more uncertain information. The underlying assumption appears to be that more uncertain data should be more difficult for the reader to examine in detail and gather specific information from, though he does not explicitly state it in this way.

Matthew McGranaghan (1993) proposed two major strategies to indicate uncertainty on maps: there can be an explicit labeling of features with text to indicate their uncertainty, much as Wright (1942) mentioned, or there can be an implicit symbology, in which "ambiguous symbols" are used to create "visual and cognitive ambiguity" for uncertain data.

Rather than being tied to a specific visual variable, McGranaghan's suggestion can be seen as a guideline for employing any such variable. A visible mark, based on uncertain data, should be drawn such that it is difficult for a user to learn any specific, detailed, and ultimately unreliable information from it. To create this ambiguity, McGranaghan suggested employing focus, as well as duplicating symbols, such that an uncertain feature is placed on a map multiple times – making it impossible to determine more than approximate spatial information. Beyond static representations, he also suggested that animation would be useful, causing uncertain features to move or change regularly.

Like MacEachren, Van der Wel *et al.* (1994) examined Bertin's variables as a starting point. They proposed, in detail, matches between different variables, the type of uncertainty being conveyed (positional accuracy, attribute accuracy, data lineage, etc.), and the level of measurement (nominal, ordinal, interval, ratio). They also suggested several non-traditional techniques: MacEachren (1992)'s advice on using focus is repeated, and the use of scale is suggested, where uncertain data would be shown at a small scale to prevent the discernment of fine details. Other proposals included rendering three-dimensional surfaces with uncertainty as height, causing uncertain information to blink on and off in a dynamic display, and a sequence of maps with the data varying within the bounds of its uncertainty.

Each of these publications has the same underlying assumption, though McGranaghan gives the clearest and most explicit articulation: more uncertain data should be more difficult to read.

Experimental Work

Technique proposals were soon followed (and in one case preceded) by experimental testing of how map readers actually performed when uncertainty was encoded in various ways. Schweizer and Goodchild (1992) asked subjects to look at maps in which uncertainty was encoded by either saturation or color value. Here, the less certain data was encoded by lighter colors or less saturated ones. Thematic data were coded by the remaining visual variable not used for uncertainty. Users were shown maps both with and without legends. They found no differences between user accuracy in map reading tasks with either symbolization scheme. In absence of a legend, the authors suggested that users generally answered on the assumption that “dark is more” (whether for uncertainty or the thematic data), though darker colors are a combination of both saturation and value. They do not state whether or not their test maps were on a light or dark background – that is, whether darker colors stood out from, or blended in to, the background.

MacEachren *et al.* (1998) conducted user testing with three different visual variables to encode data uncertainty on a choropleth map: uncertain data were shown by a change in texture, color saturation, or color hue. Uncertain data were given less vivid colors, obscured by a diagonal pattern overlay, or given a shift in hue from a purple-green scheme to a red-blue one. The former two both involve making uncertain data more difficult to extract – the less vivid colors stand out less on the white-background test maps, and the diagonal pattern overlay on the choropleths makes the underlying choropleth color less clear. The authors show an awareness of the notion of graphical ambiguity leading to cognitive ambiguity when they refer back to the variables of focus and saturation as having “the potential to suggest uncertainty or lack of precise knowledge.” The authors also tested the use of a map pair, in which a map of data uncertainty was juxtaposed with a map of the thematic data. User performance on some map reading tasks (estimating values and identifying clusters at different map scales) was worse for the color hue and color saturation techniques than the map pair or texture, though for other tasks no difference was noted.

Leitner and Buttenfield (2000) tested the variables of saturation, color value, and texture to indicate uncertainty on white background maps. Their work is the only published experiment in which each variable is tested in both directions. That is to say, the authors did not just test a scheme in which (taking saturation as an example) more certain data were given more vivid colors, but also one in which more certain data were given less vivid colors. Others before them have operated under the assumption, stated or otherwise, that more certain data should be more easily visible and stand out from the background more, and their publication appears to be the only one which tests whether or not this leads to better map reading outcomes. They concluded, based on assessments of map reader speed and accuracy, that color value should be given preference over texture in encoding uncertainty, and texture over saturation. The variables, in their opinion, should be used such that the more certain data are represented by a lighter color, a finer texture, or a more saturated color. Their guideline on the usage of value is particularly interesting, as their experiment was conducted using white-background maps. Representing more certain data by lighter colors therefore means representing more certain data with symbols that stand out less – are more ambiguous.

Edwards and Nelson (2001) experimentally tested four different techniques using proportional symbol maps. Two visual variables were tested: focus and color value. Uncertain data were given a lighter color against a white map background, or had their symbols made to be more transparent, allowing the background to show through more. Both of these experimental decisions are based upon the unstated assumption that less certain data should be more ambiguous. Two other methods were also tested, the legend statement (a verbal warning to the reader describing data uncertainty) and the reliability diagram, which is akin to the map pair described above. Those maps relying on the visual variables outperformed the other two techniques in subject accuracy and confidence in map reading tasks. Likewise, in such tasks, maps using focus to indicate uncertainty led to more accurate and confident user responses than those using color value.

Aerts *et al.* (2003), examined color value and color hue as variables to indicate uncertainty in a map of projected urban growth. In applying color value, more uncertain data were given a lighter color against the white map background. For hue, more certain data were marked in blue and uncertain in red, in a simple binary scheme. The authors also tested user preferences for showing uncertainty by a map pair and one with a sequential presentation (in which they could toggle freely between a map of the data and a map of uncertainty). Users preferred the map pair over toggling, and, for the group looking at the map pair, color value was slightly preferred over the color hue scheme.

The Need for Testing Graphical Ambiguity

McGranaghan's (1993) idea of "graphical ambiguity" is employed implicitly by many of the above researchers, who choose to make their more uncertain data stand out less on the map, by giving it a lighter color on white paper, for example. Few, however, have explicitly stated the role of this principle in their research. Edwards and Nelson (2001) make one of the only explicit references to this idea, in discussing their results with the focus variable: "More certain data gets the graphic punch, at the expense of the less certain data, so much so that perhaps it becomes a more effective means of displaying the two data sets [thematic data and uncertainty] in tandem." MacEachren *et al.* (1998), also writing while looking back at the results of their experiment, remark that, "the visually integral color scheme tested here would have the general result of preventing readers of a presentation-oriented atlas from seeing any clusters that contained a high proportion of enumeration units for which data were not reliable."

But if a symbol is to stand out more on a map, it must stand out *from something* – its background. As the background changes, a symbol can blend in more or less. An understanding of a symbol's visual context is critical to creating graphical ambiguity, if this is indeed our goal. However, the concept of graphical ambiguity itself has not yet been tested formally as a means of conveying uncertainty information. Previous experiments focus only on specific visual variables and not necessarily on their relation to their context, drawing universal conclusions on symbol schemes – Leitner and Buttenfield (2000), for example, give the guidance that more vivid colors should be used for more certain data. They do not state whether or not this depends on on the background color of the map, but make the tacit assumption, as do many of the other authors above, that visual context is not a critical factor in user interpretations of uncertainty symbols. Their suggestion may still be valid guidance – if map readers are examining symbols without considering their visual context and contrast with their background. However, if users are instead interpreting uncertainty symbols based on an assumption that less certain data is more graphically ambiguous, then surroundings must be accounted for when choosing how to apply specific visual variables. This thesis will test the validity of the assumption that visual context does not influence the reading of uncertainty symbols. Results will determine if there is a need for a broader, context-based guidance on symbol schemes for uncertainty, or whether the existing literature can be taken as is.

Only Leitner and Buttenfield (2000)'s work appears to provide evidence concerning the effect of context on uncertainty symbol interpretation, as they tested visual variables in the direction of "more certain is less distinct from the background" as well as "more certain is more distinct from the background." They found mixed results – users were more accurate in reading maps that had more certain data coded as graphically ambiguous (lighter colors on a white background) and as less graphically ambiguous (more vivid colors).

Visual Detection

Visual variables do not take into consideration how symbols relate to their context, and cannot by themselves be used to describe McGranaghan's "graphical ambiguity." Some other way of considering how symbols are perceived, and how they interact with their background, is needed. One way of grounding this concept of "standing out more" or "ambiguity" in physical, measurable phenomena is to consider the process of *detection*. Detection occurs when an observer is aware of the presence of an object in the visual field. It is distinct from such tasks as recognition or resolution, which require further processing of the visual signal. It is purely a physical phenomenon related to the quantity (and differences in) light striking the retina (Riggs, 1965). Lamar *et al.* (1947) have shown that the ability to detect an object decreases as the brightness contrast between the object and the field decreases. They have also shown that the required brightness contrast to detect an object on a field decreases as the object increases in size, though the relationship is not strictly linear. Thus, at the same brightness contrast, the detectability of an object improves as it gets larger. The concept of detectability provides an approach for testing the context-dependent concept of graphical ambiguity.

Methods

McGranaghan's suggestion of graphical ambiguity as an appropriate method for indicating uncertainty is in need of testing. The visual property of detectability allows a simple proxy for the idea of graphical ambiguity.

This thesis posits two hypotheses:

1) *In absence of a legend to guide them, map readers will tend to interpret more detectable symbols as representing more certain data, rather than less certain data:* as McGranaghan proposed, they will make a connection between uncertainty and graphical ambiguity.

2) *Readers will more accurately extract uncertainty information from maps that encode more certain data by more detectable symbols than they will from maps that encode more certain data by less detectable symbols*: if hypothesis one is true, then a map which uses graphical ambiguity to indicate increased, rather than decreased, certainty will cause map reading errors.

Method of Observation

The two hypotheses assert what map readers may or may not do. Therefore, the most suitable test for confirming these hypotheses was to observe map readers and determine if their behaviour in fact matches these assertions. A multiple-choice survey provides a simple and effective tool for making such observations. It is also a widely-accepted method in cartographic research; each of the several experimental papers described in the review section above employed a survey tool in which human subjects were shown maps and asked a planned series of questions. The hypotheses posited by this thesis are relatively simple and would be unsuitable to a more interactive observation of human subjects, such as a focus group or personal interviews – this thesis does not seek to analyze complex motivations or experiences, only simple patterns in behaviour. Thus a survey was chosen as the most suitable method for addressing the hypotheses.

In order to address the first hypothesis, subjects were shown maps and asked to compare pairs of point symbols in absence of a legend. Their answers demonstrate whether they consistently interpret the more detectable symbol of the pair as representing more or less certain data. For the second hypotheses, subjects were shown maps with legends and asked questions which required them to extract uncertainty values and compare symbol pairs to determine which is more uncertain. As a legend was provided, there was a correct answer to each question, and their accuracy could be measured.

Survey Subjects

To collect observations, a web-based survey instrument was created in Adobe Flash. It was felt that a web survey would reduce the burden on the subject (they may take it any time, anywhere, need not mail in forms, need not travel to a controlled laboratory location) and encourage greater participation. The survey was designed to take less than ten minutes to complete, again to increase the rate of response. Subjects were not compensated for their participation, as this would be impractical to carry out on a web survey.

The survey was publicized on web forums frequented primarily by cartographers and GIS specialists, in order to ensure that the responding population had a basic minimum of cartographic literacy. Hypothesis two involves testing map reading errors. By ensuring a familiarity with maps in the test population, map reading errors cannot be attributed to an inexperience reading maps generally, but can be more confidently assigned to the choice of symbols on a particular test map. Geodata professionals were also felt to be more likely to be familiar with and interested in cartographic research, and so more likely to respond with their participation.

General Survey Design

Subjects were presented with an interactive survey interface, of which an example is seen in Figure 3. On the left side of the display is a map, and on the right, a series of questions requiring the subject to interpret that map. The map depicts a fictional landscape and point data set, to avoid potentially biasing any subjects based on prior knowledge of an area. Asking subjects to examine the point symbols and comment on which was more "uncertain" without describing the type of uncertainty or to what data the certainty applied would have been a poor reflection of actual map reading situations. Therefore, to make the survey exercise similar to real-world map usage, in which a map reader is likely to see a data set *and* its associated uncertainty, a thematic data set which had a level of uncertainty was chosen. Each point on the map is described as representing a particular type of farm, and the farm crop is encoded using the visual variable of shape. The thematic data set, while not directly relevant to the hypotheses, is necessary in that it carries some uncertainty, described to the subject as the chance that the farm type was mis-categorized by an agricultural census worker. The uncertainty is encoded by either changing the size of the symbol or the color value on a white-to-black ramp. Both of these visual variables affect the overall detectability, and therefore graphical ambiguity, of the symbol.

The uncertainty data were grouped into four classes – thus, there are only a small number of sizes and color values for the subject to distinguish between. Each point symbol was marked by a letter, to be referenced by the survey questions. Water and road features were included in order to make the test situation more like reading an actual map, rather than interpreting an abstract field of points with little connection to geography or real-world map usage.

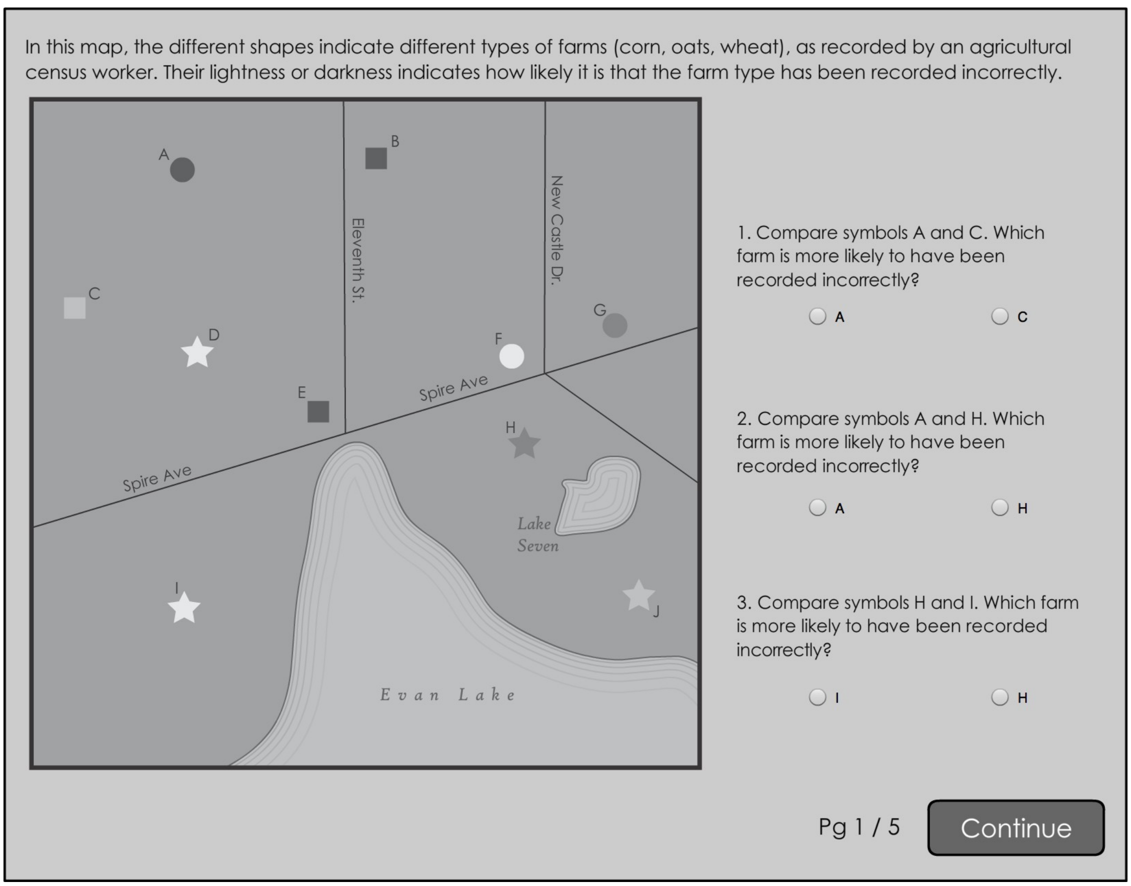

FIGURE 3 – Screenshot of example survey map and questions.

Eleven test maps were prepared in the same general style. Nine maps employed color value to encode uncertainty, and two employed size. The color value maps were divided into three groups of three maps each. Each of the three groups had a different fictional data set – a different arrangement of the point features, as well as the water and road features. Each of the three maps in a group varied only in background color – one map used a light color scheme for the land and water, one used a dark scheme, and one used a middle grey scheme.

Each of the two size maps had different data sets, with black symbols on a light grey background. The background color scheme was kept consistent for the size maps.

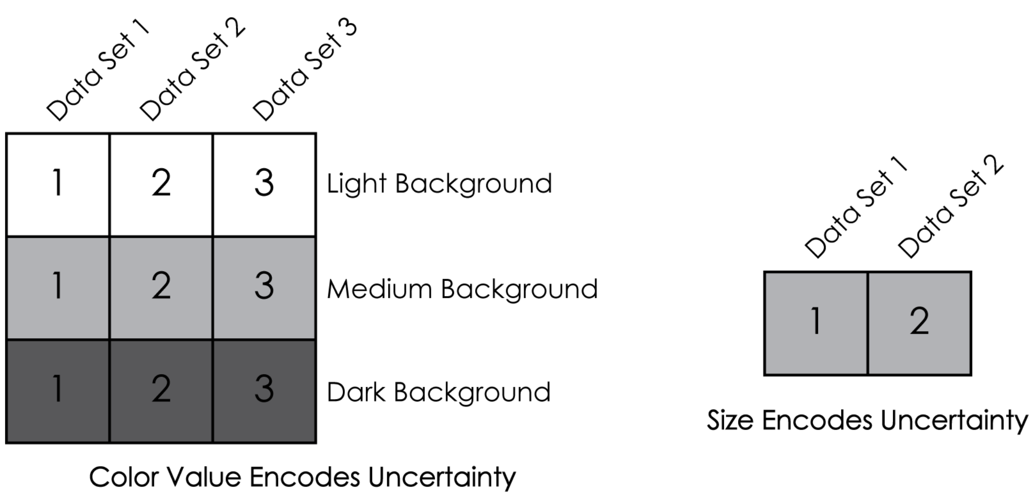

FIGURE 4 – Variations in test maps.

EXPERIMENT ONE

This experiment was designed to test the first hypothesis: *map readers will tend to interpret more detectable symbols as representing more certain data, rather than less certain data.* To examine this hypothesis, subjects were shown maps without legends and asked to perform map reading tasks so that their symbol interpretations could be determined.

Screen One

To begin, subjects were shown a randomly-chosen color value test map. A written statement at the top of the map explained the was that the map represents farm locations, with the shape indicating type, and the lightness or darkness of the symbol indicating the chance that the farm type has been mis-categorized. They were asked to compare a pair of farm symbols and answer which one is based on more certain data. No legend was provided, leaving them to make their own interpretation of the symbology, as to whether the lighter or darker symbols represent more certain data. This question is repeated twice more with new farm symbols. Each symbol pair was chosen randomly by the survey instrument. Each of the symbols in the first pair is drawn from a different class, and thus has a different color value. Another comparison pair is drawn from the remaining two classes. The last comparison pair chooses at random one symbol from the first pair and one from the second. Thus four symbols are chosen at random and formed into three question pairs, which are served in random order. This repeated testing helps to confirm a subject's interpretation of the map symbols and check for inconsistent interpretations. If, for example, a subject is chooses to interpret symbol A as being more certain than symbol B, and B as more certain than C, then they should likewise interpret A as being more certain than C. If their interpretation is, however, inconsistent – finding C to be more certain than A – it suggests that the subject has not found a clear pattern in the symbology. This experimental design is patterned after McGranaghan (1989)'s testing on choropleth interpretation. In that experiment, subjects were asked to interpret choropleth maps with various background colors, without a legend provided. Repeated comparisons of map symbol pairs determined whether or not they interpreted symbols in a "dark is more," or "light is more" direction, or whether or not they produced an inconsistent interpretation.

Screen Two

In order to help compensate for in-test learning, the first portion of the experiment was treated as a warm-up exercise – the results were not intended to be used in analysis. This was done in the belief that subjects would require some time to familiarize themselves with the survey. They need to process how to interact with the survey, to understand what is being required of them, and determine how they would like to interpret the map when no legend is given. It was believed that subjects would be less consistent in their interpretation of the map during this familiarization time, and that the results of the first few questions would not necessarily correctly reflect the subject's normal reading order of the symbols.

The second screen of the experiment functioned identically to the first. Once again, the subject was asked three times to compare symbol pairs on a color value test map. The second experimental map was drawn at random from the same pool, with the following conditions: 1) it must not be the same map as was shown on the first screen, and 2) it must have the same background color as the map shown on the first screen. This latter condition was necessary to avoid letting past experience bias a subject's interpretation. For example, a subject who is randomly shown a light-background map may choose to interpret the symbols such that darker (more detectable) symbols are more certain. If they are next shown, on the second screen, a dark-background map, they may feel a need to maintain consistency and continue interpreting darker (now, less detectable) symbols as more certain, rather than adapting to the new background color and interpreting the symbols as lighter (more detectable) equalling more certain.

Screen Three

On the third screen of the experiment, the subject was shown one of the two size-based uncertainty maps, again with three comparison questions. This portion of the experiment functioned identically to the first two screens.

Experiment Two

This experiment was designed to test the second hypothesis: *Readers will more accurately extract uncertainty information from maps that encode more certain data by more detectable symbols than they will from maps that encode more certain data by less detectable symbols.* To do so, subjects were asked to interpret maps with a legend – unlike the first experiment, there were now correct answers to each question, and subject accuracy could be measured.

Screen One

Subjects proceeded seamlessly from the conclusion of Experiment One into the first screen of Experiment Two. The subject was shown the remaining unused color value uncertainty map that matched the same background color used in Experiment One, again to avoid creating biases by changing colors. Unlike Experiment One, a legend was provided below the test map. One of two legends was selected at random: one which directed the subject to interpret darker symbols as being more certain, and one which directed them to interpret lighter symbols as being more certain. Five questions were then asked.

The first and last question drew two symbols at random from the map, and asked the subjects to compare and answer which was based on more certain data. One symbol each from the four classes was used in these two questions. The second and fourth questions asked the subject to identify what type of farm was located at a map symbol chosen by the survey instrument at random. Farm type was encoded by one of three shapes, and could be found on the legend. While this question does not directly relate to a subject's ability to extract uncertainty information off of the map, it provides a baseline assessment of their map-reading accuracy. The third question asked subjects to locate a randomly-chosen map symbol, and then identify the level of uncertainty – here described as the chance that the farm type has been miscategorized by an agricultural census worker. The correct answer could be found by examining the legend and determining which of the four classes the symbol in question matched.

Screen Two

The experiment in Screen One was repeated, this time with the remaining size-based uncertainty map and a randomly-selected legend directing the subject to either interpret larger or smaller symbols as representing more certain data.

The second experiment necessarily followed the first, to avoid biasing a subject's interpretations. Had they seen a map legend at the beginning of their survey, they may have been more likely to answer the legend-less questions based on their recollections of the legend they had seen previously, rather than engaging in a free interpretation of the symbols.

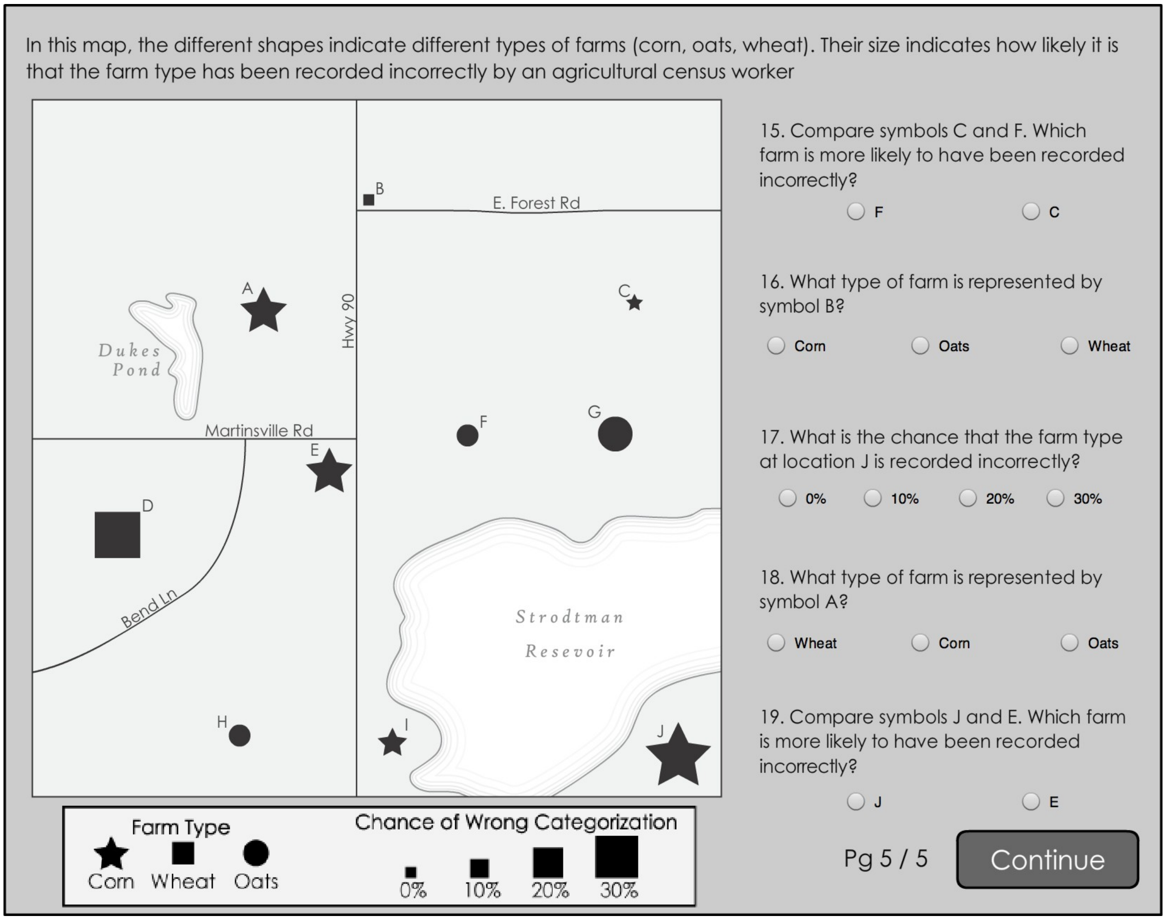

FIGURE 5 – Example of map and questions from Experiment Two, Screen Two.

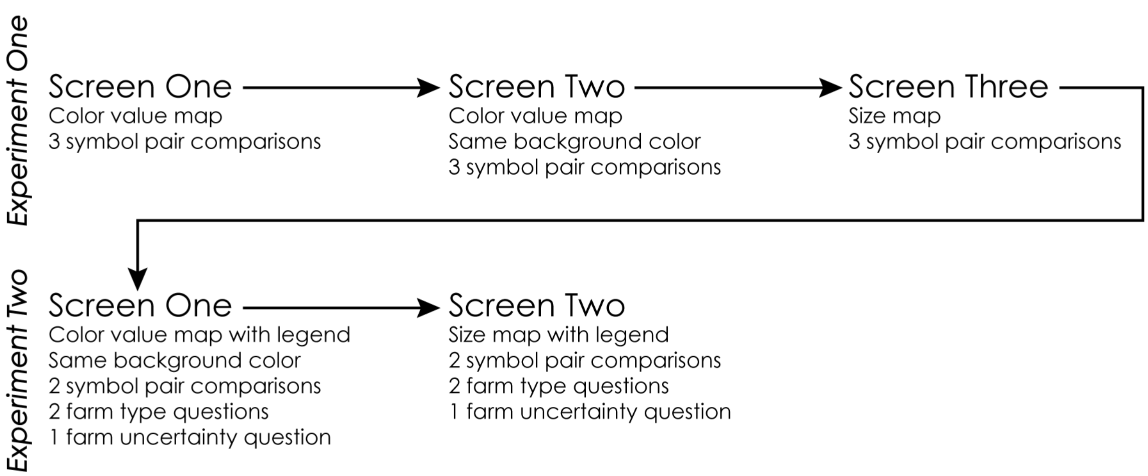

FIGURE 6 – General experimental scheme.

Results

Eight-seven responses were received. Of these, six were discarded, as the subject had answered no questions. Four responses were partial, and the remaining seventy-seven were complete.

Experiment One

Color Value Maps

As was planned initially, subject responses to the first screen of Experiment One were not considered, as the subjects were assumed to be familiarizing themselves with the interface during this time. Subject responses in Screen Two were used to assess the subject's reading scheme for uncertainty encoded by color value in symbols. As subjects were asked to perform three symbol comparisons, it was possible to determine their reading order of symbols. Three outcomes were possible: 1) the subject always interpreted the darker symbol of the pair as representing more certain data; 2) the subject always interpreted the lighter symbol of the pair as more certain; 3) the subject had a mixed interpretation, in which sometimes the lighter symbol was viewed as the more certain of the pair, and sometimes the darker.

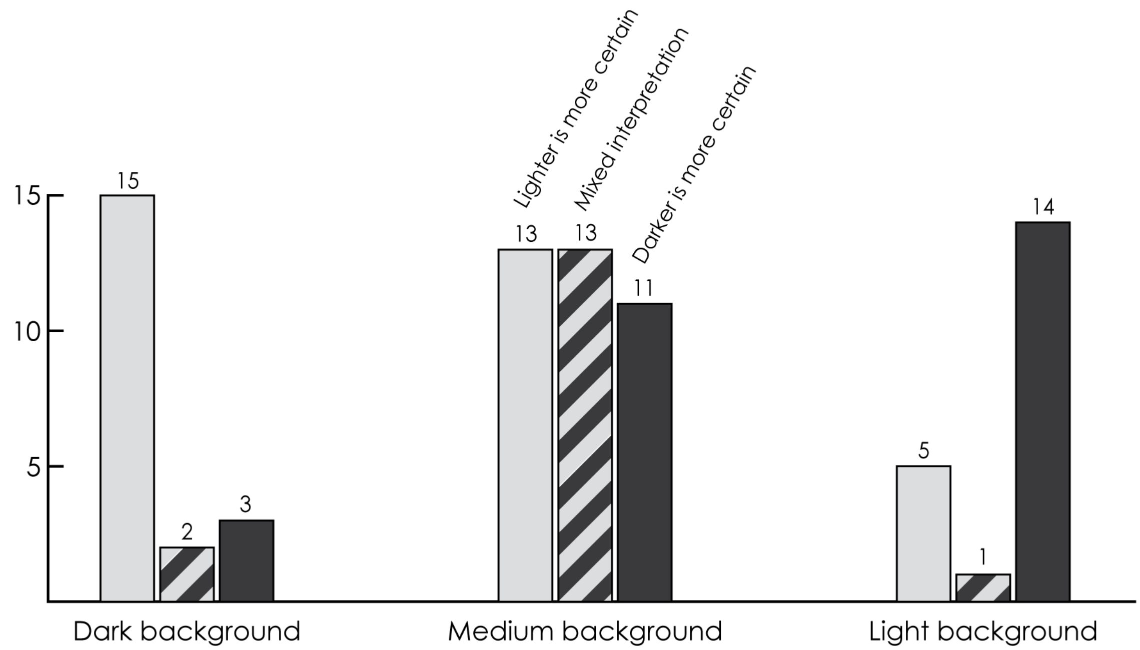

FIGURE 7 – Results of Experiment One, Screen Two.

The results of this portion of the experiment suggest that subjects' interpretation of the symbols was based upon the background color scheme of the map. On a light background map, for example, 14 out of 20 subjects interpreted the darker symbols as representing more certain data, whereas that number drops to 3 out of 20 subjects when the map background is dark. To test the statistical significance of these results, a PROB-VALUE hypothesis test was conducted, as described by Burt and Barber (1996). The test gives a <0.1% chance that the above experimental findings would result from sampling a population which was choosing at random (and thus had only a 1 in 3 chance of choosing a "lighter is more certain" interpretation). In other words, we can state with >99.9% confidence that the choice of subjects to interpret lighter symbols as representing more certain data is a reflective of a genuine choice of the population as a whole, rather than sampling error. We likewise find a similar >99.9% probability that the demonstrated subject choice of interpreting darker symbols as being more certain, against a light background, is reflective of the whole population.

Size Maps

For those maps in which size was used to encode uncertainty, subjects' symbol interpretations were likewise determined based upon the symbol comparisons performed in Screen Three, with three outcomes possible: 1) the subject always interpreted the smaller symbol of the pair as representing more certain data; 2) the subject always interpreted the larger symbol of the pair as more certain; 3) the subject had a mixed interpretation, in which sometimes the smaller symbol was viewed as the more certain of the pair, and sometimes the larger.

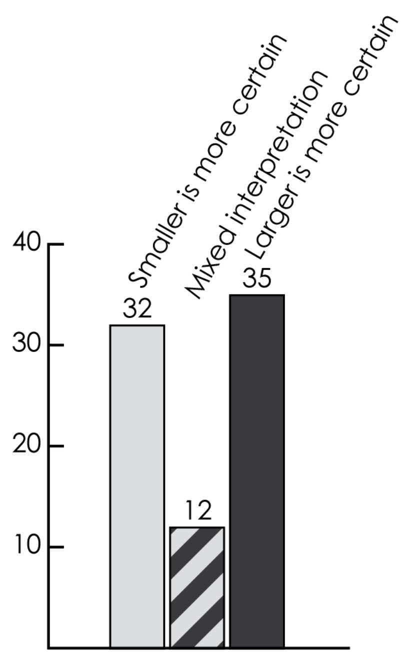

Figure 8 – Results of Experiment One, Screen Three.

The data do not reveal any clear subject consensus during this portion of the experiment. 32 of 79 subjects who participated in this section (41%) chose to interpret the smaller symbols as more certain, while 35 (44%) interpreted the larger symbols as more certain. If we set aside the subjects who had a mixed interpretation and compare the two groups which selected either smaller or larger symbols as representing more certain data, we can use the PROB-VALUE method again to determine whether or not the slightly higher number of subjects interpreting larger symbols as more certain is statistically significant. The result of the test gives a 71% probability that these results were drawn from a population which was equally likely to choose either smaller or larger symbols as more certain. The data do not demonstrate a strong relationship between detectability as realized through symbol size and the apparent certainty of the data.

A possible, though speculative, explanation is that some subjects may perceive the smaller symbols as connoting more precision. Though the fictional uncertainty for each point concerns mis-categorization of the farm type, rather than error in the placement of the point feature, a smaller mark on the page may feel more certain because it appears to be so precisely located. This sense of precision may be generalized to all uncertainty – because the subject feels more certain about the exact farm location, they also feel more confident about all data concerning that farm, including the farm type. Subjects applying this interpretation would therefore choose smaller symbols of each pair as being more certain. This interpretation does not dominate the results, however, and it the data show that other subjects are more affected by how large and therefore detectable a symbol is when assessing uncertainty.

Experiment Two

Color Value Maps – Comparing by Subject's Demonstrated Reading Order

The first screen of Experiment Two presented subjects with a legend, and asked questions which, unlike Experiment One, had correct answers. Subjects were asked to compare symbol pairs and determine which was based on more uncertain data, to use the legend to identify what farm type a symbol's shape represented, and to use the legend again to determine the exact uncertainty value for a specific point symbol. The legend, randomly selected, directed subjects either to interpret darker or lighter symbols as more certain.

The questions in Experiment One allowed a determination of how each subject interpreted the color value symbols for uncertainty: whether darker meant more certain, lighter mean more certain, or they had a mixed interpretation. Subjects in the first two categories can be put into two groups for Experiment Two: those whose previously demonstrated symbol interpretation matches the scheme given by the legend, and those whose doesn't.

Subjects who were asked to compare symbol pairs and answer which was based on more uncertain data gave correct answers to 98% of the questions when the map's symbol scheme (as dictated by the legend) matched the subject's demonstrated interpretation, as opposed to 81% correct answers when the legend did not (see Figure 9 and Table 1).

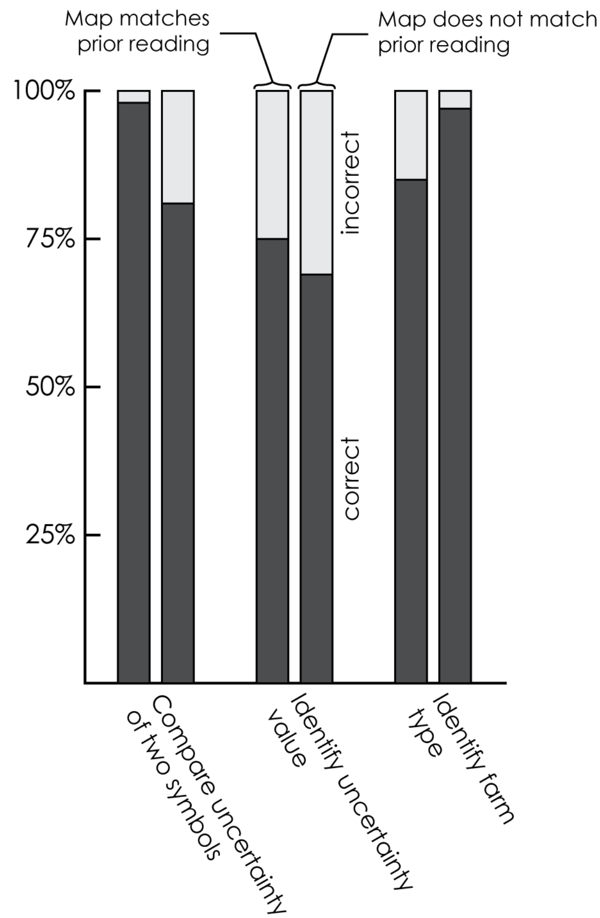

FIGURE 9 – Subject responses to color value based on whether or not map symbol scheme matches subjects' demonstrated map reading scheme.

Question Type	Map Symbol Scheme	Answers	Correct	% Correct	Population Difference in % Correct *95% Confidence Interval*
Compare uncertainty of two symbols	Matches user preference	48	47	98%	8% to 27%
	Does not match	77	62	81%	
Identify uncertainty value	Matches user preference	24	18	75%	-17% to 29%
	Does not match	39	27	69%	
Identify farm type	Matches user preference	48	41	85%	-23% to -2%
	Does not match	78	76	97%	

TABLE 1

To examine the significance of these results, the bootstrapping method, as described in Burt and Barber (1996) was used to determine if the experimental findings were due to chance or reflect a true improvement in accuracy when the symbol scheme matches the subject's previously demonstrated reading. The assumption was made that each answer to a symbol comparison question was an independent event – that is, though subjects answered two symbol comparison questions, their answers to the first question had no influence on the second. Therefore, all subject answers to each of the two questions could be pooled, as is done in Table 1 and Figure 9. The results of each of the two groups – those whose demonstrated reading was matched by the map symbol scheme, and those whose demonstrated reading was not – were resampled 25,000 times. For each bootstrap sample, the difference in the proportion of correct answers between the two groups was calculated (p_1 – p_2). The resulting distribution of differences can be used to generate a confidence interval. For this thesis, a confidence interval of 95% was selected. The results of the bootstrap test give a 95% confidence interval of 8% to 27%. That is to say, there is a 95% chance that map reader accuracy in answering symbol comparison questions increases by anywhere from 8% to 27% when the map symbol scheme matches their demonstrated reading. There is a real and demonstrable improvement.

Subjects were also asked to consult the legend and identify the uncertainty for a specific point symbol. The 95% confidence interval for the difference between the two groups is -17% to 29%. Since the interval crosses zero, we cannot say whether subject accuracy increases, decreases, or is unaffected. At the 95% confidence level, the data do not support a clear improvement or decline in map reader accuracy from a change in whether or not the symbol scheme matches their previously demonstrated choices. A possible explanation for the lack of effect is found in the fact that this question did not require the subject to compare symbols. Subjects did not have to decide how to rank symbols, with one being more than the other, and so their interpretation of how the symbols related to each other had no bearing on their answer.

When the map scheme matched the subject's earlier choice for reading the symbols, accuracy in identifying farm types (85% correct) was lower than it was when the map scheme did match (97% correct). The bootstrapped 95% confidence interval is -23% to -2%. No obvious explanation is apparent for these results, as the symbology for farm type did not change between the two groups. No significant difference for farm type questions is seem in the remaining results below.

Color Value Maps – Comparing by Detectability

Hypothesis Two specifically proposed that subject accuracy in answering questions would increase when more detectable symbols represented more certain data. The legend selected by the survey instrument can be considered as directing subjects either to interpret more detectable symbols as more certain (darker on a light background / lighter on a dark background), or the opposite. Results from subjects who were, for consistency, given a medium grey background map are not considered here, as neither of the two possible legend schemes (either progressively darker or lighter being treated as more certain) translate into the notion of "more/less detectable is more certain" in that visual context.

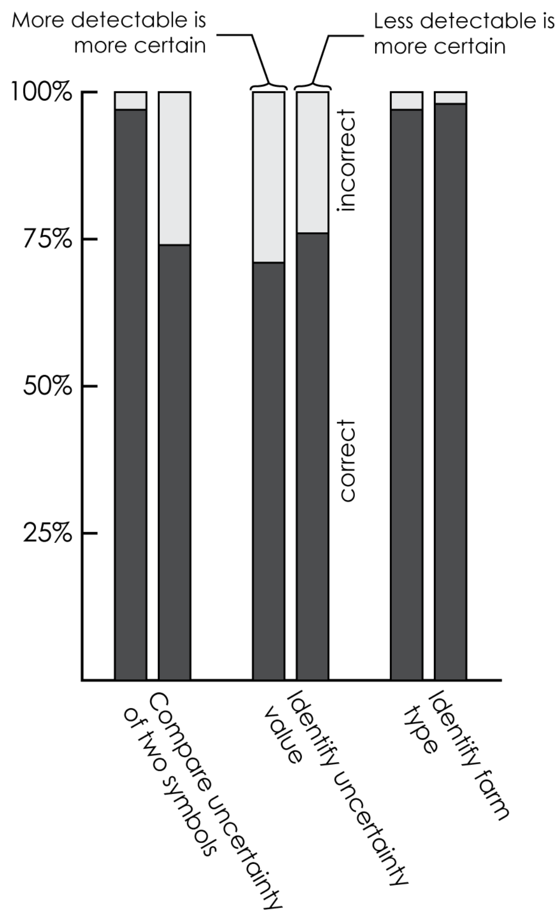

FIGURE 10 – Subject responses to color value based on whether or not more detectable symbols represented more certain data.

Question Type	Map Symbol Scheme	Answers	Correct	% Correct	Population Difference in % Correct *95% Confidence Interval*
Compare uncertainty of two symbols	More certain more detectable	33	32	97%	10% to 37%
	Less certain more detectable	50	37	74%	
Identify uncertainty value	More certain more detectable	17	12	71%	-33% to 20%
	spans 0	25	19	76%	
Identify farm type	More certain more detectable	34	33	97%	-9% to 6%
	Less certain more detectable	50	49	98%	

TABLE 2

Subjects who were asked to compare symbol pairs and answer which was based on more uncertain data gave correct answers to 97% of the questions when the map's symbol scheme (as dictated by the legend) encoded more certain data by more detectable symbols, as opposed to 74% correct answers when the less detectable symbols were used for more certain data. The confidence interval for the difference (10% to 37%) demonstrate that, when encoding uncertainty with color value, subjects are more likely to correctly compare uncertainty symbols when more certain data are represented by more detectable symbols. The results from Experiment One, and the discussion above on subject performance when the legend matched their previously demonstrated interpretation provides an explanation. When examining color value, subjects interpreted more detectable symbols as representing more certain data. Subjects were also more likely to answer questions correctly when the map matched their earlier interpretation. Therefore, a map in which more certain data is shown by more detectable symbols will match most subjects' earlier interpretations and yield more correct answers.

As above, accuracy was not affected by the symbol scheme when subjects were asked to consult the legend to derive specific values for uncertainty; the confidence interval for the difference, -33% to 20%, spans zero. Again, this may well be because the subject's chosen interpretation scheme is not called upon for this mental task. The question does not concern how symbols relate to each other or what the general pattern of symbol interpretation is, only requiring that the subject match a symbol on the map with one in the legend.

Size Maps – Comparing by Subject's Demonstrated Reading Order

The results of the questions from Screen Two of Experiment Two, in which subjects were presented with maps where size encoded uncertainty, allows a similar analysis as was done with the color value maps. Subjects were presented with a legend stating either that larger, or smaller, symbols should be interpreted as more certain. This was compared with a subject's earlier interpretation from Screen Three of Experiment One, and subjects were grouped by whether or not their demonstrated interpretation matched that of the legend.

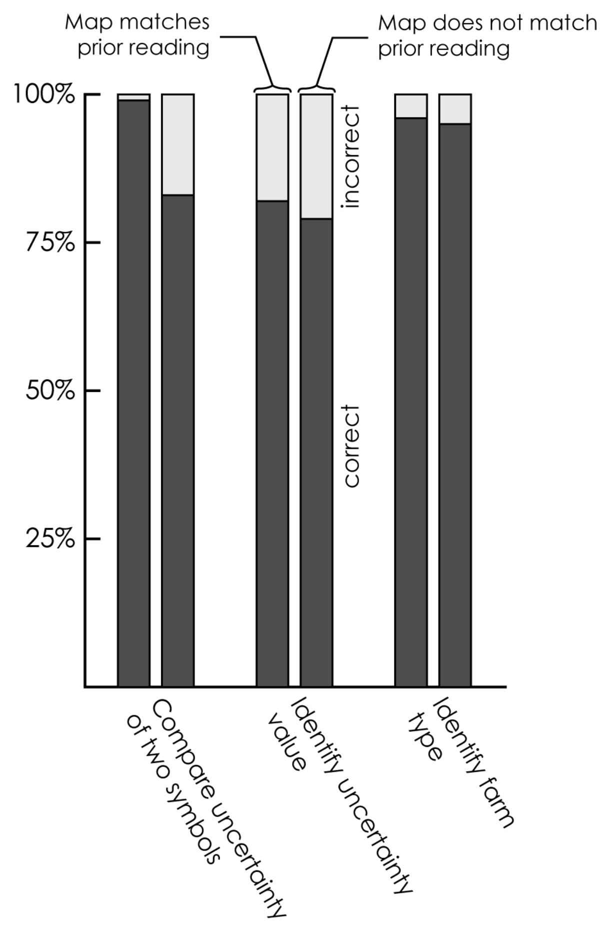

FIGURE 11 – Subject responses to size based on whether or not map symbol scheme matches subjects' previously demonstrated reading order.

Question Type	Map Symbol Scheme	Answers	Correct	% Correct	Population Difference in % Correct *95% Confidence Interval*
Compare uncertainty of two symbols	Matches user preference	68	67	99%	6% to 26%
	Does not match	66	55	83%	
Identify uncertainty value	Matches user preference	34	28	82%	-14% to 22%
	Does not match	33	26	79%	
Identify farm type	Matches user preference	69	66	96%	-7% to 8%
	Does not match	66	63	95%	

Table 3

As was found with color value, matching the subject's previous interpretation produces more correct responses to questions in which uncertainty levels must be compared. When the correct map reading matched the subject's previous map reading, 99% of responses were correct, whereas that number declined to 83% when the legend did not match. The 95% confidence interval for the difference is 6% to 26%, confirming that there is an increase in accuracy for those subjects shown maps that matched their demonstrated reading order.

Once again, performance in using the legend to extract a particular uncertainty value for a point did not differ significantly between the two groups. The confidence interval spans zero (-14% to 22%). They two groups also performed similarly in identifying farm type; the 95% confidence interval ranges from -7% to 8%.

Size Maps – Comparing by Detectability

In order to specifically test Hypothesis Two, subject responses were divided into two groups: those who saw maps in which more certain data were represented by larger (more detectable) symbols, and those who saw maps with smaller symbols representing more certain data.

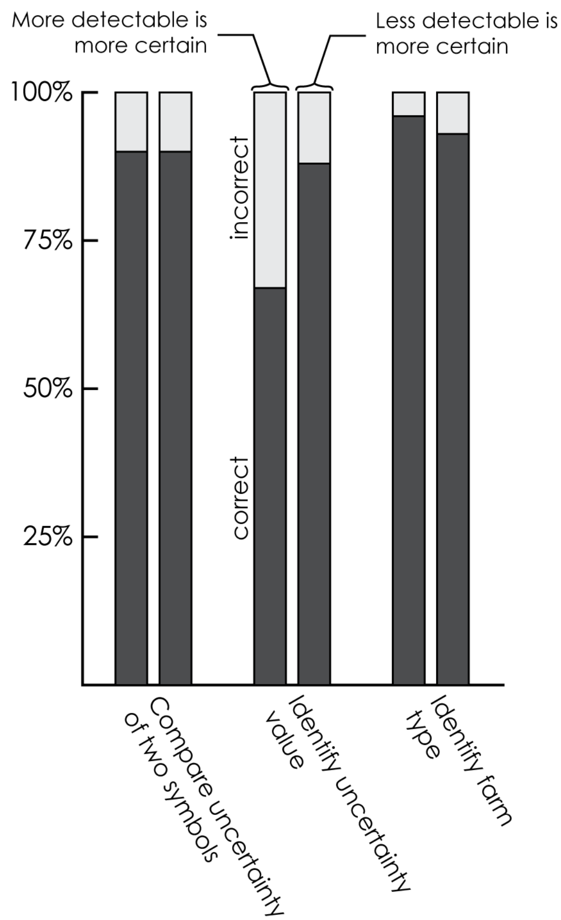

FIGURE 12 – Subject responses to size based on whether or not more detectable symbols represented more certain data.

Question Type	Map Symbol Scheme	Answers	Correct	% Correct	Population Difference in % Correct *95% Confidence Interval*
Compare uncertainty of two symbols	More certain more detectable	91	82	90%	-9% to 10%
	Less certain more detectable	68	61	90%	
Identify uncertainty value	More certain more detectable	46	31	67%	-38% to -3%
	Less certain more detectable	34	30	88%	
Identify farm type	More certain more detectable	92	88	96%	-4% to 11%
	Less certain more detectable	68	63	93%	

TABLE 4

When comparing symbols to determine which of the pair was based on more accurate data, no difference in subject accuracy was noted between the groups. Each group of subjects had a 90% accuracy in responses, and the 95% confidence interval spans -9% to 10%; since zero is included in the interval, we cannot state with confidence that there is in fact any difference between the two groups. The results of Experiment One suggest that subjects did not have a tendency to interpret either smaller or larger symbols as representing more certain data. Unlike the color value maps, where a symbol scheme that encodes more detectable as more certain was very likely to match a subject's previously demonstrated reading order, no such conjunction is possible here. Encoding more certain data with larger symbols is approximately as likely to match a subject's earlier reading order as it is to conflict with it. So, no change in accuracy when switching between larger (more detectable) or smaller (less detectable) symbols for more certain data is noted in this portion of the experiment.

The data show that subject accuracy for identifying the uncertainty value at a specific point was affected by the switch from larger to smaller symbols representing more certain data. The 95% confidence interval for the difference, -38% to -3%, shows that subjects were more accurate when less certain data were shown by larger, more detectable symbols. An obvious explanation for this result is lacking.

Conclusions

This thesis posited two hypotheses:

1) In absence of a legend to guide them, map readers will tend to interpret more detectable symbols as representing more certain data, rather than less certain data.

2) Readers will more accurately extract uncertainty information from maps that encode more certain data by more detectable symbols than they will from maps that encode certainty with the opposite scheme (less certain is more detectable).

The results of the first experiment suggest that, without the guidance of a legend, subjects establish clear and consistent schemes for interpreting uncertainty symbols. When uncertainty is encoded by color value, subjects tend to interpret those symbols with the highest contrast with the background (more detectable) as representing more certain data. When size is employed, however, subjects do not demonstrate a clear consensus for interpreting the symbols. The first hypothesis is therefore not proven – subjects do not always interpret more detectable symbols as representing more certain data. Only when the detectability is the result of changing contrast, rather than changing size, does this appear to hold true.

Subjects were asked to carry out two different map reading tasks as part of testing the second hypothesis: they compared symbol pairs to determine which was based on the more uncertain data, and they extracted specific uncertainty values for individual points. Their ability to perform the latter task did not appear to be affected by whether the more detectable symbols represented more or less certain data. Their ability to accurately compare symbols was unaffected when size was used to vary detectability, but it was affected when color value was utilized. Given the verbal construction of the second hypothesis, it is not clearly proven – subjects were in some cases able to more accurately extract information, in other cases not.

That subjects' ability to find a specific value for a point was unaffected by the symbol scheme, but their ability to make comparisons was, is explained by the fact that the former task requires use of the legend, while the latter does not. When comparing symbols on the map to gauge which is based on more certain data, the subject may err in recalling the correct interpretation of the symbols – substituting their own assumptions for what the map legend directs. If the legend were placed more prominently on the map, such that it was always in the subject's field of vision, or if the subject were made to memorize the legend beforehand, accuracy would possibly be less dependent upon what particular symbol scheme was used.

McGranaghan's notion of graphical ambiguity as a useful concept for communicating uncertainty appears sound, when graphical ambiguity is created by employing color value to generate varying levels of brightness contrast between a symbol and its surroundings. Subjects clearly rely on the visual context of a symbol to interpret it, and their responses suggested that they generally assumed the less detectable symbols, more ambiguous symbols to be those based on less certain data. The broad guidance on usage of certain visual variables by MacEachren (1992) and Leitner and Buttenfield (2000) needs to be placed in context – their results rely on the assumption of a light-colored map background, and may not hold when the map color scheme changes.

While using color value to generate ambiguity had the expected effect on subjects, usage of size did not. Smaller symbols were not associated with uncertainty any more than larger symbols. This may be because a smaller symbol conveys a sense of precision, rather than ambiguity, to some map readers.

The results demonstrate that when using color value to encode uncertainty, the visual context of a map symbol is important. Subjects tend to interpret the more detectable symbols as representing more certain data. An awareness of this tendency is important to clear communication. In a task where the legend was not immediately required, subjects were more likely to make map reading errors when the "correct" reading of the map did not match their assumptions.

Future Directions

The results of this experiment suggest new questions for future research:

- How much of a symbol's visual context affects a map reader's decisions? Are readers affected by how much a symbols stands out from its immediate surroundings, or from the color scheme of the entire map?
- To what extent are map readers' tendencies in interpreting the color value scheme dependent on their belief that the changes in color value represent uncertainty, rather than some other type of data?
- Do map readers rely on visual context for interpreting uncertainty using other visual variables, such as saturation or focus?
- How are other map reading tasks affected by mismatches between the map symbol scheme and the user's demonstrated interpretation? Can users correctly identify broad geographic trends, locate clusters, or make map-based decisions?

More work is clearly needed in uncertainty research. This thesis has taken a step in testing McGranaghan's suggestion of graphical ambiguity, but such a broad concept can be realized in may different ways than the ones employed here. Further examination of the importance of visual context in interpreting symbols for uncertainty will provide clearer, more robust guidance to cartographers who choose to communicate this type of information in their work.

Works Cited

Aerts, J. C. J. H., K. C. Clarke, and A. D. Keuper. 2003. Testing popular visualization techniques for representing model uncertainty. *Cartography and Geographic Information Science* 30(3): 249-61.

Bertin, J. 1983. *Semiology of graphics: Diagrams, networks, maps*. Translated from 1967 French edition by W. Berg, ed. Madison, WI: University of Wisconsin Press.

Burt, J., and G. Barber. 1996. *Elementary Statistics for Geographers*. New York: Guilford Press.

Edwards, L.D., and E. S. Nelson. 2001. Visualizing data certainty: A case study using graduated circle maps. *Cartographic Perspectives* 38:19-36.

Lamar, E. S., S. Hecht, S. Shlaer, and C. D. Hendley. 1947. Size, shape, and contrast in detection of targets by daylight vision. *Journal of the Optical Society of America*. 37(7): 531-45.

Leitner, M., and B. P. Buttenfield. 2000. Guidelines for the display of attribute certainty. *Cartography and Geographic Information Science* 27(1): 3-14.

MacEachren, A. M. 1992. Visualizing uncertain information. *Cartographic Perspectives* 13:10-19.

MacEachren, A. M., C. A. Brewer, and L. W. Pickle. 1998. Visualizing georeferenced data: Representing reliability of health statistics. *Environment and Planning: A* 30: 1547-61.

MacEachren, A. M. A. Robinson, S. Hopper, S. Gardner, R. Murray, M. Gahegan, and E. Hetzler. 2005. Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science* 32(3): 139-60.

McGranaghan, M. 1989. Ordering choropleth map symbols: the effect of background. *The American Cartographer* 16(4): 279-85.

McGranaghan, M. 1993. A cartographic view of spatial data quality. *Cartographica* 30 (2& 3): 8-19.

Morrison, J. L. 1974. A theoretical framework for cartographic generalization with emphasis on the process of symbolization. *International Yearbook of Cartography* 14: 115-127.

Riggs, L. 1965. Visual acuity. In: Graham, C. H. (ed), *Vision and Visual Perception.* New York: John Wiley and Sons. pp 321-49.

Roth, R. 2007. The role of mapping typological uncertainty in decision making. Unpublished Master's thesis, University of Wisconsin – Madison.

Schweizer, D. M. and M. F. Goodchild. 1992. Data quality and choropleth maps: An experiment with the use of color. In: *Proceedings, GIS/LIS '92*, San Jose, California. Washington, D.C.: ACSM and ASPRS. pp 686-99.

van del Wel, F. J. M., R. M. Hootsman, and F. Ormeling. 1994. Visualization of data quality. In: A. M. MacEachren and D. R. F. Taylor (eds), *Visualization in modern cartography.* London: Pergamon. pp 313-32.

Wright, J. K. 1942. Map makers are human: comments on the subjective in maps. *The Geographical Review* 32(4): 527-44, as ctd in McGranaghan 1993, above.