

The impact of user expertise on geographic risk assessment under uncertain conditions

Robert E. Roth | reroth@psu.edu
GeoVISTA Center, Department of Geography
The Pennsylvania State University

ABSTRACT: This research addressed the impact that expertise has on the use of uncertain geographic information for assessing the flooding risk of sites in the landscape. An online survey was developed and administered to fifty-six participants to collect initial insight into possible differences between experts and novices. In the survey, participants were required to assess the risk of an identified site relative to three delineations of the floodplain, each carrying a different degree of certainty, and then report their perceived difficulty in making the assessment and their confidence in the assessment. The responses were examined across three definitions of expertise (education/training, work experience, and self-reported) and in two realms of expertise (domain and map use). Analysis found that user confidence is affected by expertise in both domain and map use expertise, but that it is domain expertise that most impacts the risk assessment itself and map use expertise that most impacts the perceived difficulty of the risk assessment.

KEYWORDS: representing uncertainty, visualizing uncertainty, expertise, risk assessment, user-centered design, floodplain mapping.

Introduction

This research speaks to two current trends within GIScience. A first trend is the recent advance in geospatial technologies, particularly web-based applications and location-based services, which have improved the availability, affordability, and pervasiveness of geographic information. Such developments have empowered the general public with access to geographic information, and tools for representing and visualizing this information, once available only to professionals and domain experts (Rød et al., 2001; Couclelis, 2003; Wood, 2003). As this trend continues to mature, it can be expected that the general public will become increasingly active in acquiring and utilizing geographic information to support risk assessment and decision making. A second trend is the growing consensus among GIScientists that uncertainty is inherent in all geospatial datasets and “not just a flaw to be excised” (Couclelis, 2003, 166). Efforts concerning uncertainty are no longer focused solely upon reduction or elimination of uncertainties, but are rather investigating the nature of uncertainty throughout the analytical process, the forms in which it may be present in geospatial data and geographic information, and the ways in which cartographic representations and visualizations of uncertain information can be made both useful and usable for end users (Deitrick and Edsall, 2008).

The overlap of these two trends in GIScience presents the potentially disastrous situation of an untrained, inexperienced novice using representations and visualizations of geographic information that is less than certain to conduct risk assessments and make decisions that hold significant, real-world consequences. Further complicating this matter, the widespread availability of numerous data sources of the same geographic phenomenon produces a common condition of uncertainty due to lack of data agreement. While much work has been conducted to understand how uncertainty influences processes such as information assembly, risk assessment, and decision making, only a limited subset of these studies have taken into consideration the expertise of the participant within the domain at hand or with geographic information and

cartographic representations (Evans, 1997; Kobus et al., 2001; Aerts et al., 2003). This research examines the differences between experts and novices in (1) geographic risk assessments completed under uncertain conditions, (2) perceived assessment difficulty of using these representations, and (3) assessment confidence when using these representations. An important subcomponent of this research is determining if the type of expertise held by a user influences the amount of difference between experts and novices in risk assessment, perceived assessment difficulty, and assessment confidence.

To provide initial insight into the impact of expertise on these three variables, a map-based online survey was developed and administered to fifty-six participants. The following section summarizes several studies on the impact that expertise has on risk assessment predicted under uncertain conditions. In the third section, details about the case study domain of floodplain mapping are provided and the risk assessment tasks using discrete visual analog scales (DVAS) are explained. The results of the survey are reported and discussed in the fourth section and closing remarks are offered in the final section.

Literature Review

Uncertainty is present in all geospatial data, and therefore maps displaying this data, because of the inability to perfectly reconcile representations of the landscape with the actual reality of the landscape. This research defines *uncertainty* as “a measure of the user’s understanding of the difference between the contents of a dataset, and the real [geographic] phenomena that the data are believed to represent” after Longley et al. (2005, 128); the often conflated terms ambiguity, error, quality, and reliability are assumed to be part of (although not synonymous with) this larger definition of uncertainty but are not described in more detail. Defining uncertainty in this manner is perhaps controversial, as the user of the map plays an important role in its degree of certainty, suggesting that the certainty of the map is in part reliant upon external factors. However, when keeping in mind that the map is a tool that supports risk assessment, and decision making based upon these assessments, it is important to acknowledge the human component in the assessment or decision. Thus, a user-centered perspective of uncertainty is a “context-dependent concept, dependent on both the individual and the situation of the data’s creation and use” (Deitrick and Edsall, 2008, 279). This research focuses upon a single aspect of the map user important to the interpretation and use of uncertain representations: his or her level of the expertise.

At least three studies have looked at the differences between experts and novices when using representations of uncertain geographic information, with contradicting results. Evans (1997) examined the utility of several different methods for representing the uncertainty of land use/land cover (LULC) data (display of only certain pixels via map filtering, integrated symbolization using hue for LULC class and saturation for certainty, and animated flickering between the data and its certainty on a single map). Experts were defined as participants with either university-level training or professional experience in map use and interpretation. Forty-four experts and twenty-two novices completed a survey consisting of uncertainty assessment and area estimation tasks based on the aforementioned LULC representations. Evans (1997) discovered that, regardless of expertise level, a majority of map users understood the presented uncertainty information and found it helpful for completing the tasks. Further, there was no significant difference between experts and novices in response accuracy or the estimation of overall map certainty. Interestingly, experts spent more time viewing the maps than novices, meaning it took experts longer to complete the tasks than novices with no increase in accuracy.

Aerts et al. (2003) examined two methods for visualizing the uncertainty of a 2050 urban growth projection for Santa Barbara, CA using the SLEUTH simulation (side-by-side comparison of the data and its certainty on separate maps and toggling between the data and its certainty on a single map). Participants self-reported expertise in urban planning, decision support, map visualization,

or GIS; those that did not select one of these were considered novices. Thirty-seven experts and twenty-nine novices used the visualizations to estimate the rate of growth projected by the SLEUTH model and to make decisions about the growth in relation to the uncertainty of the model and other background information layers. Like Evans (1997), Aerts et al. (2003) concluded that a majority of the participants, regardless of expertise, could understand and use the uncertainty information and that inclusion of the information improved spatial decision making. Although some differences existed in preference of the two visualization methods, they were not significant. The only significant difference between groups was on the relative preference of a bivariate color scheme for representing uncertainty, as experts responded with higher affinity for a more complex bivariate scheme than did novices (although experts, like novices, still preferred the univariate scheme overall).

Finally, the findings of Kobus et al. (2001) appear to sharply contrast those of Evans (1997) and Aerts et al. (2003). Kobus et al. (2001) recorded the impact of uncertainty representations on the speed and accuracy of tactical decisions made by military officers. Citing Klein's (1993) recognition-primed decision-making model, Kobus et al. (2001) asserts that expertise plays a fundamental role in military decision making, as an initial course of action (COA) is modified once features or patterns familiar from training and experience are identified. Experts were defined as participants with at least ninety days of command post experience. Twenty-three expert and twenty-nine novice Marine Corps officers participated in a computer simulation designed to mimic a military scenario. Movements of troops were displayed on an interactive topographic map and additional information was provided via audio and onscreen text. Kobus et al. (2001) found that the experienced officers took significantly more time to develop situational awareness under uncertain conditions than novices, but were able to select and execute a COA significantly faster than their novice counterparts once awareness of the situation was established. The Kobus et al. (2001) study suggests that expertise level does in fact influence assessments or decisions predicated upon the uncertainty information.

Methodology

An online survey, built using Adobe Flash software, was developed and administered to provide insight into the impact of user expertise on risk assessment under uncertain conditions. The primary drawback to conducting research online is that the participants are not in a controlled environment; the participant can be interrupted midway, have his or her attention split on another website or activity, and be assisted by another person during the survey. However, like the Aerts et al. (2003) study, the use of an online survey was justified because of the desire to question a large amount of experts, located across the United States, in a short period of time. The domain of floodplain mapping was chosen because of the relative simplicity of binary floodplain maps (i.e., most floodplain representations depict only two categories: inside and outside of the floodplain) and their common usage by professionals and the general public alike for the purpose of risk assessment. A link to the online quantitative survey was emailed individually to 135 potential participants and was completed by 56 participants, producing a surprisingly high survey response rate of 41.5%. Among those asked were University of Wisconsin-Madison faculty and graduate students studying GIScience or fluvial processes and private, state, and federal professionals working in GIScience and floodplain mapping. Because the survey results were stored anonymously in an external database, it is unknown how well each of these groups was represented.

The survey began by asking a set of six background questions to establish each participant's level of expertise. Expertise has been defined in at least three different ways: (1) by amount of formal education or training (e.g., Evans, 1997), (2) by level of work experience (e.g., Evans, 1997; Kobus et al., 2001), and (3) by self-reporting of personal expertise (e.g., Aerts et al., 2003). Further, it is possible that expertise in the case study domain of floodplain mapping is

fundamentally different than expertise in map use and interpretation to support risk assessment and decision making (Aerts et al., 2003), producing six categories of expertise for comparison. A single question in the background survey was structured to define participant expertise in each of these six categories. For background questions defining expertise by education/training and work experience, participants could select one of two levels of expertise (educated/trained versus uneducated/untrained, experienced versus inexperienced). For background questions defining expertise by self-reporting, participants could select one of three levels of expertise (expert, intermediate, or novice). The middle category ‘intermediate’ for the self-reporting definition was included because a tendency of participants to report themselves as ‘expert’ if the only other option available was ‘novice’ was revealed in a pilot study. Table 1 provides a break down of the 56 participants according to the six categories of expertise.

| <i>Expertise Category</i> | | <i>Expert</i> | <i>Intermediate</i> | <i>Novice</i> |
|---------------------------|--------------------|---------------|---------------------|---------------|
| domain | Education/Training | 26 | n/a | 30 |
| | Work Experience | 21 | n/a | 35 |
| | Self-Reporting | 10 | 34 | 12 |
| map use | Education/Training | 47 | n/a | 9 |
| | Work Experience | 42 | n/a | 14 |
| | Self-Reporting | 40 | 14 | 2 |

Table 1: Survey participation between experts and novices. This study adopted three definitions of expertise (education/training, work experience, and self-reporting) and identified two different realms of relevant expertise (domain and map use).

Following the background survey, the participants completed a brief training session introducing each component in the layout of the online survey and demonstrating how each component should be used for completion of the survey. Once finished, participants had the option of reviewing the training session again or continuing to the survey itself. Figure 1 provides a screenshot of the main portion of the survey. The layout of the main portion of the online survey consisted of three components: (1) a map displaying three artificial delineations of the floodplain relative to a river along with one of three marked sites (Figure 1A), (2) a legend describing the symbology on the maps (Figure 1B), and (3) a triplet of questions concerning the current map-legend pair (Figure 1C). Each component is described briefly in the following subsections.

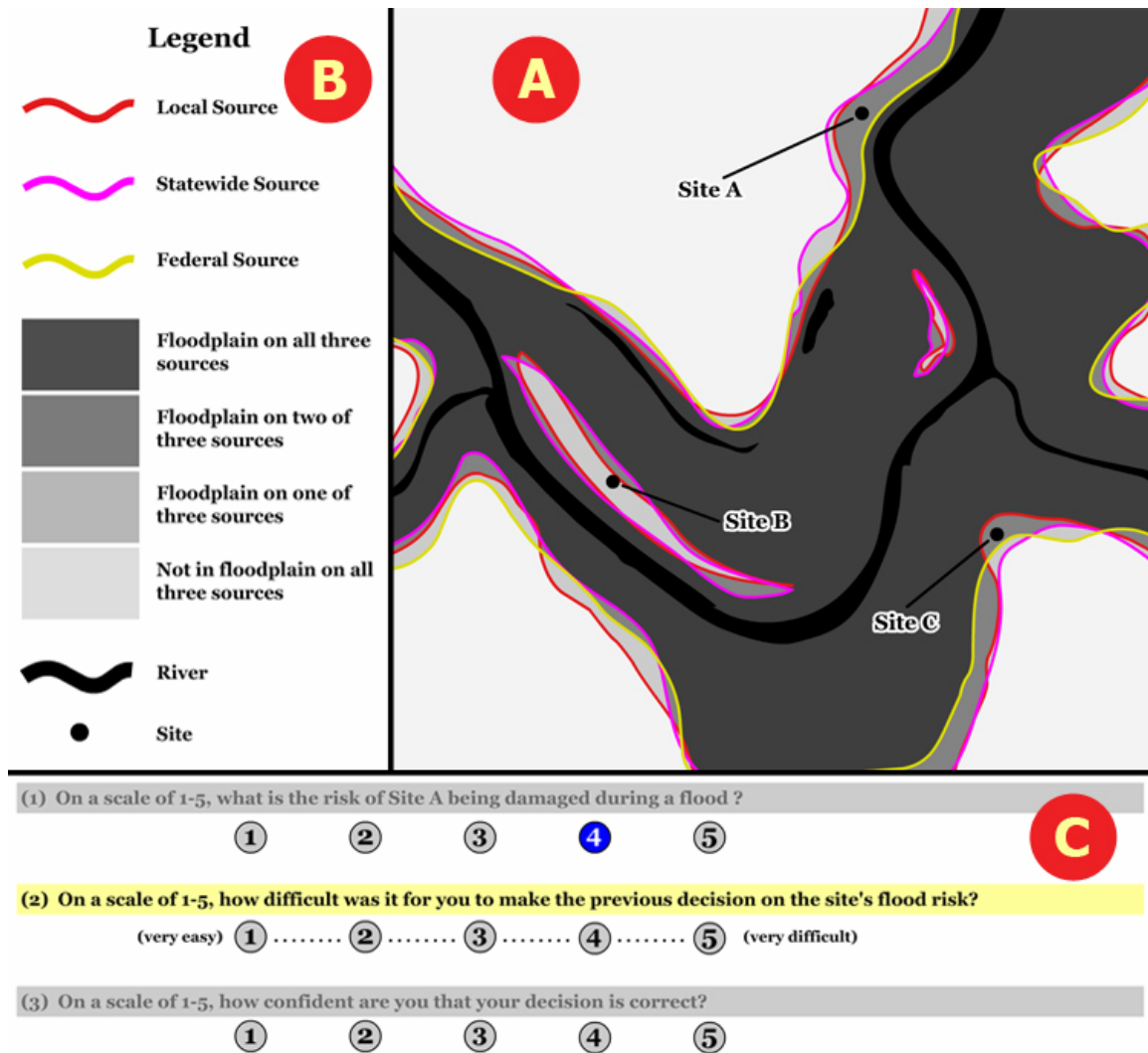


Figure 1: The layout for the main portion of the online survey. (A) The map component of the survey displayed three artificial delineations of the floodplain relative to the river and one marked site for reference by the question component. The above example shows all three of the sites used in the survey for reference, but the actual survey displayed only a single site at a time. (B) The legend component of the survey described the symbology present on the maps. The upper portion of the legend suggested an ordinal level of certainty in a given uncertainty category for each floodplain delineation. (C) The question component of the survey consisted of three questions using discrete visual analog scales (DVAS). Responses to these three questions defined the variables risk assessment, perceived assessment difficulty, and assessment confidence respectively. The question component was the only interactive portion of the survey.

The Map Component of the Survey

The central component of the survey was a static map displaying the location of a river, three artificial delineations of the floodplain relative to this river, and a single location or site of interest. Provision of three separate floodplain delineations was a powerful and real-world map-use scenario for testing the influence of expertise on risk assessment, as both experts and novices commonly use several, possibly contradicting delineations of the floodplain for assessing the flooding risk of a particular site in the landscape, such as an existing home or potential construction site (e.g., Lulloff, 1994). Such a situation produces a condition of uncertainty for risk assessment and decision making due to lack of data agreement. To conform to common landscape patterns produced by fluvial processes, the river and floodplain delineations in the survey were based upon floodplain data for Albany, Oregon, located along the Willamette River (Catlin,

2002). The floodplain delineation was modified to produce four sets of three artificial floodplain variants (a total of twelve delineations), each set corresponding to a different sub-component or category of uncertainty. Inclusion of multiple categories of uncertainty matched the broad definition of uncertainty offered in the previous section and also ensured that survey responses reflect each participant's full comprehension of uncertainty rather than his or her familiarity with a single category of uncertainty. Generation of the floodplain delineation variants for each uncertainty category was informed by supplemental geographic information about the local terrain and discussion of the nature of the uncertainty category provided by Thomson et al. (2005) and MacEachren et al. (2005). The decision to use artificial delineations reduced the influence of prior participant knowledge, ensuring that participants have never seen the floodplain delineations upon which the risk assessments were predicated.

The four categories of uncertainty represented in the map component of the survey were (1) credibility, (2) currency, (3) precision/resolution, and (4) subjectivity (as defined by MacEachren et al., 2005). Floodplain delineation variants for MacEachren et al.'s (2005) uncertainty categories of accuracy/error and completeness were also included in the survey, but survey responses were not included in the analysis because the resulting maps were either much more complex (as with the case of accuracy/error, where temporal accuracy was displayed) or much less complex (with the case of completeness, where only two variants were provided) than the maps for other uncertainty categories; such disparity in complexity produced a condition of inconsistent map interpretation difficulty, skewing the responses for maps displaying accuracy/error and completeness. The variants for a single uncertainty category were generated to produce an ordinal decrease in certainty (i.e., high certainty, medium certainty, low certainty). A definition of each uncertainty category and a description of the three floodplain delineation variants generated for the survey are provided in Table 2. The desire to represent multiple uncertainty categories in the experiment further justified the use of artificial delineations, as real data containing metadata on four categories of uncertainty was not available.

| <i>Uncertainty Category</i> | <i>Definition</i> | <i>Variation #1</i> | <i>Variation #2</i> | <i>Variation #3</i> |
|-----------------------------|--|---|---|---|
| Credibility | reliability of the information source | federal source | statewide source | local source |
| Currency | time span from occurrence through information collection/processing to use | data collected in 2005 | data collected in 1995 | data collected in 1985 |
| Precision/Resolution | exactness of measurement/estimate | high detail | intermediate detail | low detail |
| Subjectivity | the extent of human interpretation/judgment | floodplain defined by a simulated discharge level | floodplain defined by the historical record | floodplain defined by innermost terrace |

Table 2: The four uncertainty categories used in the survey, their definitions according to MacEachren et al. (2005), and a description of the three variants for each category as used in the legend component of the survey.

It was decided not to represent the certainty of the floodplain delineations directly on the map, as floodplain maps rarely provide graphic uncertainty information in an integrated display. Floodplain maps instead rely upon a single verbal statement of data certainty for the entire map extent. To communicate uncertainty type, each floodplain delineation was symbolized by a unique hue to allow for categorical discrimination. The relative certainty of the colored

delineations was explained by the legend component of the survey in a verbal statement (see the following subsection). A qualitative color scheme was used so that participants could not determine a delineation variant's certainty by the map representation alone. Such a strategy provided insight into the participants' abilities to rank the three delineations for each uncertainty category based on the verbal statements provided by the legend. Each map displayed the three variants of a single uncertainty category (i.e., the twelve artificial delineations were not mixed randomly but kept organized by associated uncertainty category).

Three sites within the map extent were chosen for flood risk assessment, as shown by Figure 1A. For each category of uncertainty used in the survey, there were three accompanying versions of the map, each displaying only one of these sites (note that Figure 1A shows all three sites for illustrative purposes). Site-A and Site-C were contained by two of the three floodplain delineation variations in all four uncertainty categories. However, Site-A was contained by the two most certain floodplain delineations in each of the four uncertainty categories while Site-C was contained by the two least certain floodplain delineations in each of the four uncertainty categories. Site-B was contained by only one floodplain delineation: the least certain in each of the four uncertainty categories. Therefore, the expected flood risk ranking (from highest flood risk to lowest flood risk) was A-C-B for all four uncertainty categories. The sites were symbolized using a large black dot and labeled using a 14pt black font with white casing for reference in the survey questions. The background of the map was represented in varying shades of gray corresponding to the amount of overlap or agreement among the three floodplain depictions, reinforcing the condition of uncertainty to the participants. The river and surrounding oxbow lakes, unmodified from the Catlin (2002) data source, were symbolized in black to reflect the highest category in the grayscale representation of floodplain certainty.

Several controls other than the use of artificial floodplain delineations were implemented in the map component to avoid the use of previous knowledge of the area and to discourage learning throughout the course of the survey. To disguise the geographic location of the dataset, all signifiers were removed from the map, including information on roads, houses, geographic markers, and place names. To reduce the effect of learning during the survey, each newly loaded map was randomly rotated around its center. The order of maps was also randomized for each participant following Leitner and Bittenfield (2000), making it unlikely that two participants saw the same order of maps. Finally, the question component of the survey was disabled until several seconds after the map was fully loaded and visible to prevent input of previous answers and accidental clicks.

The Legend Component of the Survey

The legend component of the online survey described each symbol included in the map component. The top portion of the legend provided a verbal statement defining each floodplain delineation displayed in the map component; the verbal statements for each variant included in the survey are provided in Table 2. Beard and Mackaness (1993) contend that there are three uncertainty assessment tasks of increasing difficulty: (1) notification (presence of uncertainty), (2) identification (what kind of uncertainty is present and its relative amount), and (3) quantification (the magnitude of uncertainty). To avoid blending quantification assessments with identification assessments in the survey, categories commonly reported at the ratio level (e.g., precision/resolution) were worded in the legend to match uncertainty categories commonly reported at the ordinal or categorical level (e.g., credibility). The middle portion of the legend explained the grey-scale shading used for the amount of overlap or agreement among the three floodplain delineations. The bottom portion of the legend explained the symbology used for the river itself and the sites of interest.

Because understanding the legend is vital to understanding the map, and therefore correctly assessing the flooding risk of a location in the landscape, several precautions were taken to ensure that the legend was read before interpreting the map and responding to the survey questions. First, the importance of the legend was stressed in the initial training session and the participants were

instructed to first analyze the newly loaded legend before examining the map. Secondly, as a result of a pilot survey, the legend was placed at the top, left corner of the screen. Eye-movements studies have shown that Western readers tend to process a page from the top, left corner to the bottom, right corner, similar to the way text is read in a book (Slocum et al., 2005). Placing the legend at the top, left corner put it in the highest priority location on the screen and ensured that the participant did not need to scroll the browser to view it. Similarly, the most important information, the verbal statements explaining the three delineation variants, was placed at the top of the legend component for immediate reading. Third, a loading transition was programmed to darken the screen before loading a new map-legend pair to signal that the legend required reference. Finally, as with the map component, the question component of the survey was disabled until several seconds after the legend was fully loaded and visible, preventing response before the new legend was loaded and visible.

The Question Component of the Survey

The final component of the survey was a set of three questions. Each question was presented as a *discrete visual analog scale* (DVAS). A DVAS is similar to the more commonly known Likert scale and was first introduced as a method for patients to report their levels of pain or discomfort (DeVellis, 2003). The DVAS and Likert scale both rely upon a horizontally-aligned visual scale of evenly-spaced integers to produce quantifiable metrics; once presented with a question or statement, the subject selects the integer along this scale that best matches his or her assessment, attitude, or opinion. Unlike the Likert scale, which technically must diverge from a central or neutral response and label the meaning of each individual integer, the more generic DVAS allows for sequential steps along a continuum with no middle-point and requires labeling of only the poles of the continuum (e.g., low risk-high risk, easy-difficult, confident-not confident).

Participants were required to answer three DVAS questions for each of the twelve map-legend pairs. The first question required the participant to assess the flood risk of the site on a five-level DVAS ('1' being safely located and '5' being insecurely located). The variable *risk assessment* was defined as the response to this first DVAS. The second and third questions for each map-pair were follow-up questions for the initial risk assessment. The second question asked the participant to rank the difficulty of the risk assessment on a five-level DVAS ('1' being an easy siting decision and '5' being a difficult siting decision). The variable *assessment difficulty* was defined as the response to this second question. Although the three versions of each uncertainty map were intended to create a varying degree of assessment difficulty throughout the survey much like the Leitner and Battenfield (2000) study, this second question allowed for the recording of a perceived assessment difficulty independent of the investigator-defined assessment difficulty. The third question asked the participant to rank his or her confidence in the risk assessment on a five-level DVAS ('1' being least confident that the decision was correct and '5' being most confident that the decision was correct). The variable *assessment confidence* was defined as the response to this third question (after Evans, 1997). Figure 1C illustrates each DVAS for these three variables. The current DVAS is highlighted and labeled to prompt the participant (e.g., perceived assessment difficulty in Figure 1C) and answers to previous statements are highlighted in blue for reference (e.g., a response of '4' to the risk assessment DVAS in Figure 1C).

Results and Discussion

Nonparametric testing was required for statistical analysis of DVAS responses because (1) the random variable recorded by a DVAS is not continuous (i.e., measurement of the variable is on the ordinal level of measurement rather than interval/ratio) and (2) DVAS responses typically are not normally distributed (McGrew and Monroe, 2000). Application of parametric testing under such conditions increases the likelihood that a significant finding is spurious, making nonparametric testing a more robust approach. Two nonparametric statistics were applied to the

survey responses: the Mann-Whitney U and the Kruskal-Wallis H . The Mann-Whitney and Kruskal-Wallis statistics differ from their parametric alternatives (two-sample Student t and multiple sample ANOVA tests, respectively) in that they are applied to the rankings of the data within the total sample space rather than the raw data itself. To calculate the rankings, the dataset was ordered from smallest to largest, independent of expertise grouping, and a ranking from 1 to n was assigned to each response. Tied values were resolved by using the average of their ranks, following Hollander and Wolfe (1999). The rankings were aggregated by expertise level to produce an overall sum for each grouping; the nonparametric statistics were calculated using these summations. Because all expertise groups were of sufficient size once pooled by identified site ($n=10$ for Mann-Whitney and $n=5$ for Kruskal-Wallis), a normal approximation (z) was applied for hypothesis testing of the Mann-Whitney U statistic and a chi-square (χ^2) approximation was applied for hypothesis testing of the Kruskal-Wallis H statistic. Nonparametric testing using these approximations contains 95% of the explanatory power as their parametric counterparts (Aczel and Sounderpandian, 2006). For all hypothesis tests, the null hypothesis was that there was no significant difference between/among the expertise groupings and the alternative hypothesis was that there was a significant difference between/among the expertise groupings (i.e., a two-tailed hypothesis test).

There were three components to the statistical analysis: (1) the category of expertise, (2) the site displayed on the map (each exhibiting a different degree of risk), and (3) the variable measured by the DVAS. Statistical analysis was completed across each of the six categories of expertise. The Mann-Whitney test was applied for the education/training and work experience definitions of expertise because the biographical questions allowed for differentiation between just two groups of expertise, while the Kruskal-Wallis test was applied for the self-reporting definition of expertise because the biographical questions allowed for three groupings of expertise. For each category of expertise, the analysis was applied four times: once pooling all responses in the survey together and once for the responses to each of the three identified sites (pooling maps showing the same site but different types of uncertainty together). Initial examination of all DVAS responses pooled together provided an overview of the interaction for a single category of expertise. However, it was also important to examine the responses for each site individually because the varying level of risk associated with each site could possibly produce a clouded signal for the pooled hypothesis tests. Uncertainty type was not accounted for because each site had the same relative amount of certainty regardless of the type of uncertainty described in the legend (i.e., Site-A exhibited the highest flooding risk on all four types of uncertainty) and examination at the individual question level (i.e., a specific site/uncertainty type combination) produced sample sizes too small for reliable analysis. Finally, the statistical analysis was applied to each of the three variables collected by the survey: (1) risk assessment, (2) perceived assessment difficulty, and (3) assessment confidence. Discussion of the results is organized around these three variables. In sum, sixty-four hypothesis tests were conducted. Care was taken during interpretation of the test results, as it likely to receive one or several spurious significant findings running such a large number of hypothesis tests.

The first DVAS variable analyzed was risk assessment. Z -scores or chi-square values, depending on the applied nonparametric test, and p -values for each hypothesis test completed on the risk assessment variable are provided in Table 3. The p -values of hypothesis tests that reject the null at $\alpha=0.05$ are bolded and italicized in the table. An important pattern uncovered by the survey responses was that experts or intermediates, regardless of expertise category, responded with higher risk assessments than their novice counterparts for all three sites (with the two exceptions of the map use work experience category for Site-B and the map use self-reported category for Site-C). Assuming that flooding risk assessments from experts are more accurate than those from novices, this pattern in risk assessment suggested a tendency for novices to underestimate the flooding risk of a site in the landscape. One possible explanation of this tendency is that domain experts better comprehend the implications of the uncertain conditions than domain novices and increase their risk assessment accordingly. The tendency for novice underestimation of risk

appeared especially strong when domain expertise was considered, as five of the six domain expertise hypothesis tests for Site-A (the site at most risk) and Site-B (the site at least risk) were extremely significance, with the sixth showing significance at $\alpha=0.10$. It was possible that none of the hypothesis tests returned significance for Site-C under domain expertise because the site exhibited a medium amount of relative risk; some participants, regardless of expertise, appeared to interpret it as more similar to the high-risk Site-A while others interpreted it as more similar to the low-risk Site-B, producing a mixed signal for both expert and novice groupings. It was likely that only one of the three hypothesis tests applied to domain expertise when all sites were pooled together returned significance because of this mixed signal for Site-C, although it is important to note that all three tests showed significance at $\alpha=0.10$. The connection between map use expertise and risk assessment was much weaker, as only one of the twelve hypothesis tests returned significance at $\alpha=0.05$ and only two at $\alpha=0.10$. This suggested that it is a lack of expertise in the domain at hand, rather than in map use and interpretation, causing the underestimation of risk by novices. Map use expertise does not appear to affect risk assessment significantly.

| <i>Category</i> | | <i>Site-A</i> | | <i>Site-B</i> | | <i>Site-C</i> | | <i>All Sites</i> | |
|-----------------|--------------------------|------------------------|----------------|------------------------|----------------|------------------------|----------------|------------------------|----------------|
| | | <i>z/X²</i> | <i>p-value</i> | <i>z/X²</i> | <i>p-value</i> | <i>z/X²</i> | <i>p-value</i> | <i>z/X²</i> | <i>p-value</i> |
| domain | Education/Training | 11.1860 | 0.0000 | 1.7282 | 0.0894 | 0.5582 | 0.5767 | 1.6906 | 0.0909 |
| | Work Experience | 2.4096 | 0.0160 | 1.8676 | 0.0182 | 0.8593 | 0.3902 | 1.9564 | 0.0504 |
| | Self-Reporting (X^2) | 10.1220 | 0.0063 | 8.2244 | 0.0164 | 1.5298 | 0.4654 | 9.2882 | 0.0096 |
| map use | Education/Training | 0.8702 | 0.3842 | 0.7972 | 0.4253 | 2.1236 | 0.0337 | 1.7319 | 0.0833 |
| | Work Experience | 0.9869 | 0.3237 | -0.7240 | 0.4691 | 1.2286 | 0.2192 | 1.0107 | 0.3122 |
| | Self-Reporting (X^2) | 0.7314 | 0.6937 | 1.3923 | 0.4985 | 1.6459 | 0.4391 | 0.7194 | 0.6979 |

Table 3: Analysis of the variable risk assessment. Overall, experts responded with higher assessments of risk than their novice counterparts. The difference between experts and novices appears especially strong when domain expertise is considered, suggesting that is expertise in the domain at hand, and not in map use and interpretation, that affects a user's ability to correctly assess the risk of a site in the landscape.

Responses to the DVAS variable perceived assessment difficulty were then analyzed; hypothesis testing for perceived assessment difficulty is summarized in Table 4. Generally, intermediate and novice participants found the assessment tasks to be more difficult than their expert counterparts (explaining the negative z-scores), although this pattern is violated for all four hypothesis tests conducted on the education/training definition of domain expertise. The reasoning for this unintuitive violation is unclear, but may be an artifact of recruiting a large amount of current graduate students. Similar to the risk assessment variable, there is no significant difference across expertise in perceived assessment difficulty of Site-C; it is again hypothesized that this is due to its medium amount of flooding risk. The most striking finding for the perceived assessment difficulty variable is that five of the six definitions of expertise report significance when pooling all sites together. This suggested that expertise does play an important role in the perceived difficulty of completing an assessment task under uncertain conditions. Interestingly, when examining responses for Site-A and Site-B, the division between expert and novice is stronger for map use expertise (four of six tests for Site-A and Site-B returning significance) than domain

expertise (only one of six tests for Site-A and Site-B returning significance). Thus, while domain expertise significantly influences the risk assessment under uncertain conditions, it is expertise in map use and interpretation that influences the perceived difficulty in completing the assessment. Such a finding provides initial evidence that domain and map use expertise are two separate characteristics of a potential map user and need to be accounted for separately.

| <i>Category</i> | | <i>Site-A</i> | | <i>Site-B</i> | | <i>Site-C</i> | | <i>All Sites</i> | |
|-----------------|---|------------------------|----------------|------------------------|----------------|------------------------|----------------|------------------------|----------------|
| | | <i>z/X²</i> | <i>p-value</i> | <i>z/X²</i> | <i>p-value</i> | <i>z/X²</i> | <i>p-value</i> | <i>z/X²</i> | <i>p-value</i> |
| domain | Education/Training | 0.4589 | 0.6463 | 1.6538 | 0.0982 | 1.6176 | 0.1057 | 2.2051 | 0.0274 |
| | Work Experience | -1.1170 | 0.2640 | -0.3260 | 0.7444 | -0.3590 | 0.7196 | -0.9851 | 0.3246 |
| | Self-Reporting (<i>X²</i>) | 10.6270 | 0.0049 | 0.4079 | 0.8155 | 5.3699 | 0.0682 | 10.9560 | 0.0042 |
| map use | Education/Training | -2.3020 | 0.0213 | -1.5020 | 0.1331 | 0.3355 | 0.7372 | -2.0016 | 0.0453 |
| | Work Experience | -2.1540 | 0.0312 | -3.5430 | 0.0004 | -0.6750 | 0.4997 | -3.7026 | 0.0002 |
| | Self-Reporting (<i>X²</i>) | 3.6201 | 0.1636 | 10.6520 | 0.0049 | 0.8139 | 0.6657 | 10.4169 | 0.0055 |

Table 4: Analysis of the variable perceived assessment difficulty. Overall, experts responded with a lower perceived assessment difficulty than their novice counterparts. The difference between experts and novices appears especially strong when map use expertise is considered, suggesting that is expertise in map use and interpretation, and not in the domain at hand, that affects a user's ability to correct assess the risk of a site in the landscape.

The DVAS variable assessment confidence was the last to be analyzed; results of the hypothesis tests for the assessment confidence variable are provided in Table 5. Of the three DVAS variables, assessment confidence exhibited the most significant differences between experts and novices in both quantity and magnitude. Experts felt much more confident about their risk assessments than their intermediate or novice counterparts (with the lone exception of the domain education/training definition on the Site-B map). This finding is intuitive, as experts are expected to be much more comfortable assessing risk under uncertain conditions, and therefore more confident in their assessments, than novices. Five of the six definitions of expertise reported a significance difference when all sites were pooled together. Further, of the twenty-four hypothesis conducted, nineteen are significant at alpha=0.10, seventeen at alpha=0.05, and twelve at alpha=0.01. Although the pattern holds for both domain and map use experts, it appears slightly stronger for map use experts. This minor difference is explained by a relative lack of confidence by experts in their assessment of Site-B (lowest risk), reporting confidence values close to those reported by novices. Such a lack of confidence in the assessment of the site in the lowest flooding risk is not immediately intuitive. However, follow-up questioning of several domain experts revealed that this is likely an artifact of placing Site-B on an island in the river. Domain experts assessed the risk of flooding according to the position and certainty of the floodplain delineations, but tempered their confidence in the assessment due to the domain knowledge that a river island may wash away completely during a flooding episode (i.e., although the risk is small, if a flood does occur it can be catastrophic to the site).

| Category | | Site-A | | Site-B | | Site-C | | All Sites | |
|----------|--------------------------|---------|---------------|---------|---------------|---------|---------------|-----------|---------------|
| | | z/X^2 | p-value | z/X^2 | p-value | z/X^2 | p-value | z/X^2 | p-value |
| domain | Education/Training | 2.1138 | 0.0345 | -0.8590 | 0.3903 | 1.0336 | 0.3013 | 1.3178 | 0.1876 |
| | Work Experience | 3.5032 | 0.0005 | 0.4632 | 0.6432 | 2.2520 | 0.0243 | 3.5959 | 0.0003 |
| | Self-Reporting (X^2) | 17.7920 | 0.0001 | 4.9002 | 0.0863 | 12.4640 | 0.0020 | 30.3764 | 0.0001 |
| map use | Education/Training | 2.7230 | 0.0065 | 2.8352 | 0.0046 | 1.8612 | 0.0627 | 4.3968 | 0.0000 |
| | Work Experience | 2.9298 | 0.0034 | 2.6536 | 0.0080 | 1.9226 | 0.0545 | 4.4029 | 0.0000 |
| | Self-Reporting (X^2) | 6.0243 | 0.0492 | 6.5678 | 0.0375 | 4.2550 | 0.1191 | 16.2600 | 0.0003 |

Table 5: Analysis of the variable assessment confidence. Overall, experts felt more confident in their risk assessments than their novice counterparts. This pattern appears to hold up for both domain and map use expertise.

The results of the study strongly suggest that user expertise plays a significant role in all three collected variables, concurring with the findings from Kobus et al. (2001). Perhaps the most important finding of the study was that the six categories of expertise did not impact the three variables equally. Hypothesis testing suggested that an accurate risk assessment under uncertain conditions was the product of domain expertise, but ease in making this assessment under uncertain conditions was the product of map use expertise. Therefore, it appears as though domain experts better comprehend the uncertain information and are better equipped to incorporate this information into their analysis of the sites, leading to appropriately higher assessments of risk. However, map use experts cognitively process the cartographic representations and complete the risk assessments more easily, regardless if their risk assessments are accurate. Assessment confidence was the only variable of the three measured by a DVAS to exhibit significant differences for both domain and map use expertise. Except for the unusual finding for the perceived assessment difficulty variable for the domain education/training category of expertise, the way that expertise was defined (e.g, education/training, work experience, or self-reported) did not appear to influence participant response.

There are several important implications of these findings. First, a user that is both a domain and map use novice will tend to underestimate the flooding risk of a site in the landscape under uncertain conditions. However, coming with this assessment is also a large perceived difficulty and low amount of confidence. Therefore, it is likely that this user will seek the opinion of an expert if the consequences of the risk are severe, making a disaster possible but not probable. Second, the optimal risk analyst under uncertain conditions is a user that holds expertise in the domain at hand (so that an accurate risk assessment can be made) and in map use (so that the assessment is completed easily). Therefore, training programs should focus equally on providing the domain knowledge necessary for making the risk assessment (i.e., coursework on fluvial process and risk management) as well as instruction on map use and interpretation to improve the efficiency of this risk assessment (i.e., coursework on the cartographic basics). Third, perhaps the most disastrous situation comes when a map use expert, but a domain novice, attempts to complete risk assessment tasks under uncertain conditions. This user will find the assessment task to be relatively easy, and will perhaps be confident in their assessment. However, because this user is a domain novice, the assessment likely will underestimate the risk of the site. Therefore, the may make the unfortunate mistake of using the site for some sort of activity without consulting an expert, possibly leading to damage of persons and property. A domain expert, but

map use novice, may have difficulty in completing the assessment, but will assess the risk accurately, avoiding unintended consequences.

Conclusion

The concept of uncertainty recently was reframed within GIScience to include the vital role of the end user (e.g., Longley et al., 2005; Deitrick and Edsall, 2008). This research examined a single user characteristic that is of possible importance: user expertise. Research regarding the impact of user expertise on geographic risk assessment and decision making under uncertain conditions has produced contradictory results. Studies completed by Evans (1997) and Aerts et al. (2003) suggested that user expertise has little affect on the comprehension and use of uncertain information to complete map-based tasks, while a study completed by Kobus et al. (2001) suggested the opposite. A possible reason for this mismatch is the slightly different way in which the concept of expertise is defined in each of these three studies and the different tasks required of the participants. To provide further evidence to the open question in GIScience of the impact of user expertise on geographic risk assessment under uncertain conditions, an online survey in the domain of floodplain mapping was developed and proctored to fifty-six participants with varying amount of expertise across three definitions of expertise (education/training, work experience, self-reported) and in two realms of expertise (domain and map use). Using discrete visual analog scales (DVAS), three variables were collected: a flooding risk assessment of a site in the landscape relative to three floodplain delineations of varying certainty, the user's perceived difficulty in completing this risk assessment, and the user's confidence in the risk assessment. Hypothesis testing suggested that there were differences across domain expertise for risk assessment, differences across map use expertise for perceived assessment difficulty, and differences across both domain and map use expertise for assessment confidence.

So what does this mean for practicing cartographers designing the representations and visualizations of uncertainty for support of risk assessment or decision making? The key to designing useful and usable representations and visualizations of uncertainty (or of anything else) is to know the end user. This means constant consultation with domain experts to better understand the geographic phenomenon and the tasks that the representation or visualization need to support. This also means testing prototypes with potential end users at all stages of preparation and iteratively adjusting the design accordingly. However, cartographers must be aware that their representations and visualizations will be used for unintended purposes, particularly when made publically available online. This means, above all else, being honest about the certainty of the representation by explicitly symbolizing it on the map so that end user can make fully informed risk assessments and decisions (or, in the case of novices, to know that they should defer to experts).

REFERENCES

- Aczel, A.D. and Sounderpandian, J. (2006) *Complete Business Statistics*, 6th edition. New York, NY: McGraw-Hill.
- Aerts, J.C., Clarke, K.C. and Keuper, A.D. (2003) Testing Popular Visualization Techniques for Representing Model Uncertainty. *Cartography and Geographic Information Science*, 30, 3, pp. 249-261.
- Beard, K. and Mackaness, W. (1993) Visual Access to Data Quality in Geographical Information Systems. *Cartographica*, 30, 3, pp. 37-45.
- Catlin, R. (2002) Floodplain (cited map). Albany, OR: The City of Albany Planning Division.
- Couclelis, H. (2003) The Certainty of Uncertainty: GIS and the Limits of Geographic Knowledge. *Transactions in GIS*, 7, 2, pp. 165-175.
- Deitrick, S. and Edsall, R.M. (2008) Making uncertainty usable: Approaches for visualizing geographic uncertainty. In: (eds) Dodge, M., McDerby, M. and Turner, M. *Geographic visualization: Concepts, tools, applications*. London: Wiley & Sons.
- DeVellis, R.F. (2003) *Scale Development: Theory and Applications*, 2nd edition. Newbury Park, CA: Sage Publications.
- Evans, B.J. (1997) Dynamic Display of Spatial Data-Reliability: Does it Benefit the Map User? *Computers & Geosciences*, 23, 4, pp. 409-422.
- Hollander, M. and Wolfe, D.A. (1999) *Nonparametric Statistical Methods*, 2nd Edition. New York, NY: John Wiley and Sons.
- Klein, G.A. (1993) Recognition-Primed Decision (RPD) Model of Rapid Decision Making. In: (eds) Klein, G.A., Orasanu, J. and Calderwood, R. *Decision Making in Action: Models and Methods*. Norwood, NJ: Ablex.
- Kobus, D.A., Proctor, S. and Holste, S. (2001) Effects of experience and uncertainty during dynamic decision making. *International Journal of Industrial Ergonomics*, 23, pp. 275-290.

Leitner, M. and Buttenfield, B.P. (2000) Guidelines for the Display of Attribute Certainty. *Cartography and Geographic Information Science*, 27, 1, pp. 3-14.

Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W. (2005) *Geographic Information Systems and Science*, 2nd edition. Hoboken, NJ: John Wiley & Sons.

Lulloff, A.R. (1994) Winnebago County – Floodplain Redelineation Project. In: *Proceedings of Urban and Regional Information Systems Association (URISA) Annual Conference*. Milwaukee, WI, pp. 675-679.

MacEachren, A.M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M. and Hetzler, E. (2005) Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartography and Geographic Information Science*, 32, 3, pp. 139-160.

McGrew, J.C. and Monroe, C.B. (2000) *An Introduction to Statistical Problem Solving in Geography*, 2nd edition. New York, NY: McGraw-Hill.

Rød, J.K., Ormeling, F. and Elzakker, C.V. (2001) An agenda for democratising cartographic visualisation. *Norsk Geografisk Tidsskrift-Norwegian Journal of Geography*, 55, pp. 38-41.

Slocum, T.A., McMaster, R.B., Kessler, F.C. and Howard, H.H. (2005) *Thematic Cartography and Geographic Visualization*, Second. Upper Saddle River, NJ, USA: Pearson Prentice Hall.

Thomson, J., Hetzler, B., MacEachren, A., Gahegan, M. and Pavel, M. (2005) Typology for visualizing uncertainty. In: *Proceedings of Conference on Visualization and Data Analysis (part of the IS&T/SPIE Symposium on Electronic Imaging)*. San Jose, CA, pp. 146-157.

Wood, M. (2003) Some Personal Reflections on Change...The Past and Future of Cartography. *The Cartographic Journal*, 40, 2, pp. 111-115.