A qualitative approach for understanding the role of geographic information uncertainty during decision making.

Abstract:
Much recent research in GIScience is focused on developing a deep comprehension of the underlying nature of uncertainty to the end of proper uncertainty representation for informed decision-making. As it is impossible to eliminate all uncertainty from an abstracted representation, it is important to understand the involvement of uncertainty in the geographic information life cycle, the many forms that uncertainty can take, and the influence these forms have on decision-making. This paper examines the involvement and influence of geographic information uncertainty during decision-making using the case study domain of floodplain mapping. A set of focus groups composed of floodplain mapping experts was conducted to provide initial insight into the following research questions: (1) How is uncertainty involved in the creation, representation, and use of geographic information in the domain of floodplain mapping and how can this practice be improved? (2) Is the MacEachren et al. typology an appropriate categorization of the many geographic information uncertainties in the domain of floodplain mapping or are there categories that must be added, removed, or revised? (3) Which categories of uncertainty are the most influential on the decision-making process in the domain of floodplain mapping? Although the focus groups revealed that the current involvement of geographic information uncertainty is less than ideal, there was clear consensus that the MacEachren et al. typology is an appropriate categorization of geographic information uncertainty for the domain of floodplain mapping and that the categories accuracy/error, precision/resolution, and currency are the most influential on the decision-making process.

Keywords: geographic information uncertainty, uncertainty categorization, decision-making, MacEachren et al. typology, floodplain mapping

1. Introduction

> "We actually made a map of the country on the scale of a mile to a mile!" [said Mein Herr]
> "Have you used it much?" I enquired.
> "It has never been spread out, yet," said Mein Herr: "the farmers objected: they said it would cover the whole country and shut out the sunlight! So we now use the country itself, as its own map, and I assure you it does nearly as well." (Carroll 1893|1982, 726.)

The above Lewis Carroll passage, well known among cartographers and GIScientists, illustrates a central concern with the representation of geographic information. In order to be understandable and usable, a map must abstract reality, removing unnecessary or less important details while maintaining, and therefore accentuating, features of interest. As noted in the Carroll passage, a map that does not abstract carries little semiotic advantage and is no more useful than interacting with the world itself. By abstracting reality, however, information is removed that may be requisite for a clear and comprehensive understanding of geographic phenomena or processes. Tasks employing the representation or decisions informed by the representation can no longer be accomplished with complete certainty. This is the *cartographic problematic*: in order to create an abstraction of reality that makes complex geographic information understandable and usable, uncertainty is introduced into the representation (and into the knowledge constructed from the representation) as a necessary compromise.

Because uncertainty is inherent to all geographic information, and therefore decisions based on it, research in GIScience should focus on better ways to manage and use uncertainty during decision-making rather than attempting to purge it from all geographic information (Couclelis 2003; Deitrick and Edsall 2008). There are at least three challenges to coping with geographic information uncertainty: (1) determining the current involvement of uncertainty at different stages in the geographic information lifecycle (i.e., is it even collected and represented for use in a decision?), (2) identifying the many forms that uncertainty can take in this process (i.e., do decision-makers need to consider multiple categories of uncertainty and, if so, what are they?), and (3) understanding the influence these forms have on the use of geographic information (i.e., are all uncertainties considered as equal when making a decision?).

Unfortunately, there are currently few empirically-based solutions for these three challenges. Although there is a strong emphasis in the GIScience literature on developing new methods for collecting and representing uncertainty information, there is rarely a follow-up component to these studies investigating if and how the suggested techniques are applied in practice (challenge #1). On the other hand, many scholars have attempted to partition the broad concept of uncertainty and define its components (challenge #2), but with little agreement. A nine-category typology of uncertainty first offered by Thomson et al. (2005) then restated and elaborated upon by MacEachren et al. (2005) is examined in this research because it is specific to uncertainties that influence information analysts. Finally, despite the large amount of writing on GIScience uncertainty typologies, authors have only speculated on the relative influence of these categories on the decision-making process (challenge #3), instead examining decision-making influence with a broad definition of uncertainty.

This research addresses the aforementioned three challenges using the case study domain of floodplain mapping. Floodplain maps are a formalized cartographic tool used for the evaluation of flood liabilities. Urban planners use these maps to assess flood hazards before developing an area. Similarly, emergency agencies and private insurance firms use these maps to evaluate the vulnerability of features that are already in the built landscape, whose construction cannot now be avoided. Such a mapping application provides a powerful scenario in which to investigate the involvement and influence of geographic information uncertainty on the decision-making process. Decisions based upon floodplain representations hold real world significance, with millions of dollars and lives hanging in the balance. Determining techniques for successful uncertainty integration into floodplain maps supports the making of informed decisions concerning the placement and risk assessment of features in the landscape, helping both the individual owner and the construction or insurance firm.

A set of focus groups composed of floodplain mapping experts was conducted to provide initial insight into the following research questions:

(1) How is uncertainty involved in the creation, representation, and use of geographic information in the domain of floodplain mapping? How can this practice be improved? (challenge #1)

(2) Is the MacEachren et al. (2005) typology an appropriate categorization for discussing the many geographic information uncertainties in the domain of floodplain mapping? Are there categories that must be added, removed, or revised? (challenge #2) Establishing an appropriate typology is necessary prior to evaluation of the relative influence of uncertainty categories on decision-making.

(3) Which categories of uncertainty are the most influential on the decision-making process in the domain of floodplain mapping? Which are the least influential? (challenge #3)

This paper proceeds in four sections. The following section provides an overview of the background materials and literature relevant to the three challenges listed above. In the third section, the focus group methodology is described. Results of the focus groups are summarized in the fourth section and are organized to address the above three research questions. Concluding remarks are offered in the final section.

2. Background

2.1 The involvement of uncertainty in the decision-making process

Uncertainty involvement refers to the creation, representation, or use of geographic information uncertainty. Such involvement is driven by the attitudes and opinions of GIScientists, as they must acknowledge uncertainty as an important characteristic of geographic information in order to be motivated to capture and map it. Domain end users too must acknowledge uncertainty as an inherent circumstance or condition of geographically-based tasks and decisions in order to incorporate it into the completion of tasks and decisions. In this regard, GIScientists have focused upon the development of new methods for collecting and representing geographic uncertainty information. Similarly, scholars in many domains have examined how uncertainty information should be incorporated into the decision-making process. Unfortunately, it is unknown if this theory is actually put into practice.

Although there is little direct empirical evidence, Agumya and Hunter (2002) offer a helpful conceptual framework for understanding the involvement of uncertainty in the decision-making process (see Figure 1). Drawing from the discipline of risk management, Agumya and Hunter (2002) assert that most decisions under uncertain conditions are simplified to a comparison of an estimated level of risk with a threshold level of acceptable risk, producing an appropriate risk response. The Agumya and Hunter (2002) framework illustrates how the initial involvement of uncertainty in geographic information should cause uncertainty to be involved in every step in the decision-making process. Because uncertainty is involved in the input, uncertainty is also involved in the definition of risk scenarios and the estimated level of risk. Further, because the level of risk cannot be estimated with full certainty when using uncertain geographic information, uncertainty is also involved in the risk response.

Extending Couclelis (2003), I take the position that the role of the GIScientist is not to make perfect geographic information (nor do I think this is even possible), excising uncertainty from the input dataset to provide ideal decision-making conditions throughout the Agumya and Hunter (2002) model. Instead, it is the two-part role of the GIScientist to minimize the uncertainty added during information creation *and* to explicitly involve uncertainty at each stage in the geographic information lifecycle so that the domain user has a fully informed understanding of the situation.

2.2 Defining a typology of uncertainty categories

What is included under the heading of 'uncertainty' is itself uncertain. It has been argued that other broad concepts such as *ambiguity* (differing perceptions about or imperfect indicators of a geographic phenomenon), *fuzziness* (the degree to which a geographic phenomenon belongs to a particular definition, class, or other attribute), *quality* (the fitness of geographic information to a particular application), and *vagueness* (inherently inexact definition of a geographic phenomenon or its attributes) are separate from the concept of *uncertainty* (the difference between a user's understanding of the geographic phenomenon and the real geographic phenomenon) (Longley et al. 2005; Deitrick and Edsall 2008). Like Longley et al. (2005, 129), I take a "catch-all view" of the concept of geographic information uncertainty, acknowledging that such a definition is perhaps overly-inclusive and contestable. Further philosophical discussion on the nature of uncertainty is outside the scope of this research.

Uncertainty categorization references the designation, articulation, and analysis of the unique subcomponents that together constitute the multifaceted concept of uncertainty. Several typologies of uncertainty specific to geographic information are offered in the literature after Sinton's (1978) recognition that uncertainty has multiple categories that need to be treated differently during measurement and representation. MacEachren (1992) argues that uncertainty derives from two distinguishable sources, (1) accuracy and (2) precision, for each of Sinton's (1978) three components of geographic information (space, time, and attribute). A second typology offered in the literature adopts the Spatial Data Transfer Standard (SDTS), listing five categories of uncertainty: (1) positional accuracy, (2) attribute accuracy, (3) logical consistency,

(4) completeness, and (5) lineage (Buttenfield 1993). A third typology, offered by Zhu (2005), includes a different set of five uncertainty categories: (1) accuracy, (2) precision, (3) resolution, (4) consistency, and (5) completeness.

Finally, MacEachren et al. (2005) offer a typology based upon the components suggested by Thomson et al. (2005). Nine categories are identified by MacEachren et al. (2005): (1) accuracy/error, (2) completeness, (3) consistency, (4) credibility, (5) currency, (6) interrelatedness, (7) lineage, (8) precision/resolution, and (9) subjectivity. The MacEachren et al. (2005) typology is more applicable than its predecessors because: (1) the typology is based upon a listing of uncertainties identified by information analysts as influential to their decision-making; (2) the typology is described in a tiered approach (i.e., the nine identified categories are only at the top level), allowing for interoperability across domains and a clearer discussion on how other uncertainties identified in the literature fit into the typology; (3) its categories are statistically grounded, allowing for quantifiable metrics; and (4) its categories are relevant to each of Sinton's (1978) three components of geographic information. A summary of the MacEachren et al. (2005) typology with examples when applied to the domain of floodplain mapping is provided in Roth (2009).

2.3 The influence of uncertainty on the decision-making process

Uncertainty influence references the impact that awareness of geographic information uncertainty has on the decision-making process. Impact has been interpreted in several ways, including the accuracy of the decision outcome, the speed it takes to arrive at a decision outcome, the perceived difficulty in arriving at the decision outcome, and the decision-maker's confidence in the decision outcome. Previous studies have examined whether the representation of uncertainty influences decision-making positively, by clarifying the actual geographic information, or negatively, by cluttering and confusing it (Harrower 2003). It was initially believed that uncertainty information acted much like any other type of geographic information during multivariate representation in that its inclusion only made the map more difficult to use during decision-making. Such an assumption led Beard and Mackaness (1993) and McGranaghan (1993) to caution against the representation of uncertainty, warning that it may be necessary to avoid its symbolization so that the primary information does not become clouded and unusable. Leitner and Buttenfield (2000) provided dissenting evidence against this hypothesis, demonstrating that the integration of uncertainty representations actually decreases the time it takes to make decisions by clarifying the underlying geographic information and by increasing the confidence the decision-makers has in his or her decision.

The focus groups extend this discussion by examining the relative influence of each category of uncertainty in the MacEachren et al. (2005) typology. Here, uncertainty influence is the importance of recording and disclosing a particular category of uncertainty for making fully informed decisions, as articulated by the decision-makers themselves. If a category of uncertainty is perceived as influential, its communication to decision-makers has the potential to significantly alter both the way that the information representations are used during decision-making and the decision outcomes. Those categories perceived to be more influential should be given a higher priority for involvement in both the creation and representation of geographic information. Conversely, a category of uncertainty perceived to be relatively less influential on decision-making than other categories, or entirely non-influential, is a lesser priority for involvement in geographic information creation and representation.

3. Methodology

A set of focus groups was conducted to provide answers to the research questions posed in the introductory section. A qualitative approach was taken to move beyond the *outcomes* of the decision-making process and instead to learn how uncertainties in geographic information influence the *process* itself. A *focus group* is a method of research that utilizes group discussion to solicit ideas and feedback about a concept or product (Morgan 1998). Focus groups allow for participants to expound upon the reasoning behind a decision outcome. As Harrower et al. (2000, 283) note, “The point of a focus group is not to reach consensus, or to solve a specific problem, but rather to understand how and why people respond to something.” I am primarily concerned with understanding how and why uncertainty is involved in and influential to the decision-making process; quantitative methods that record only the outcome of a decision cannot support investigation of such a research question.

The group setting of this method provides researchers with the full range of opinions, experiences, and knowledge bases regarding a research topic and their amount of agreement among participants. Although the role of the focus group moderator is fundamental to sound focus group research, data collected during focus groups are considered less influenced by investigator bias in relation to traditional, closed-ended interviews due to the freeform interaction allotted participants during discussion (Monmonier and Gluck 1994). This interaction produces a synergistic effect as participants discuss, clarify, and append each other’s contributions, rather than simply responding to a moderator (Cameron 2005). Because of group synergy, it is argued that focus groups provide in-depth insights concerning a research topic much more quickly than quantitative methods or other qualitative methods (Stewart and Shamdasani 1990). However, the insights generated from focus groups are less generalizable beyond a specific case study than quantitative methods because of the smaller required sample size and the inherent subjectivity when interpreting results (Cameron 2005). Because of this, I am only able to comment on the involvement and influence of geographic information uncertainty during decision-making in the domain of floodplain mapping, leaving broader claims to other domains speculative.

Following Morgan (1998), the focus group research was organized in four stages: (1) planning, (2) recruiting, (3) moderating, and (4) analyzing /reporting the results. During the planning stage, the goals of the focus group are determined and the appropriate test subjects are identified. In the recruiting stage, a call for participation is circulated and qualified subjects are selected. During the moderating stage, the focus group protocol is developed, the moderator is trained, and the focus groups are conducted. Finally, the data generated from the focus groups is transcribed, codified, and interpreted during the analyzing and reporting stage. The following four subsections describe the planning, recruiting, moderating, and analysis respectively. The entirety of the subsequent results section is reserved for reporting.

3.1 Planning

The purpose of the planning stage of focus group research is to establish project goals and to identify appropriate test subjects. The focus group research goals match the three research questions bulleted in the introductory section. Many focus group studies are interested in attaining only expert feedback. In these situations, "participants are chosen on the basis of their experience related to the research topic" (Cameron 2005, 121). Choosing subjects based upon their ability to provide in-depth, relevant information is called *purposive sampling* and is considered a superior approach to random or convenient sampling in most research contexts (Patton 1990). Because the three research questions required participants to have a large experience base in floodplain mapping from which to draw commentary, only domain experts were included. While this decision significantly decreased the pool of available participants, causing recruiting issues (see the next subsection), it was deemed appropriate because those not holding domain expertise could provide very little credible insight into the research questions.

3.2 Recruiting

One of the most difficult aspects of recruiting is finding a meeting place and time that works for all participants (Kessler 2000). This issue is particularly a concern when participation is restricted to domain experts, as it is common for a community, government unit, or institution to employ only a small group of experts. The 2007 Wisconsin Land Information Association (WLIA) conference provided an opportunity to get multiple floodplain mapping experts from across the state in the same room. The WLIA is an organization composed of professionals from governmental and private firms that deal with geospatial technologies. Domain experts for the focus groups were recruited using the gatekeeper technique, a method that utilizes a single or several key individuals (the gatekeepers) to gain access to many other individuals within an organization who fit the ideal participant description (Valentine 1997). The gatekeeper was Ann Barrett, the Executive Services Manager for the WLIA. Ann Barrett authorized the focus groups for the conference and provided a contact list of all WLIA members.

There is very little agreement in the literature about the required number of participants for a focus group study. A published review of focus group studies found that typical sessions include five to eight participants, ranging from as low as three to as high as twelve (Twohig and Putnam 2002). Based on this review, a minimum of three participants per session was set, with a preference for five to eight. An open call for participation to all WLIA members attending the conference was circulated, producing ten respondents meeting the necessary requirements for participation. Two sixty-minute focus group sessions of five participants were organized around the conference schedules of the ten participants.

Morgan (1998) suggests the practice of over-recruiting by two participants per session, as it is very common for last second cancellations or no-shows. Unfortunately, the small sample of experts did not allow for such over-recruiting and only three participants attended each session (two cancellations within 24-hours and two no-shows). While a session size of three may not work when using randomly selected participants that have spent little time thinking about the research topic, I argue that the small size was less problematic with domain experts who engage with the topic daily. Further, keeping the focus groups at a size of three was deemed more appropriate than including non- or partial-experts due to the research questions. Finally, Brown (1999) argues that the number of participants per session can be much lower when the groups are purposively sampled to be homogenous, as increased membership adds little new insight, whereas heterogeneous groupings need to be much larger to ensure that the full variation of opinion is captured.

3.3 Moderating

The first step in the moderating stage is development of the *focus group protocol*, a detailed agenda outlining focus group discussion (Monmonier and Gluck 1994). Five types of questions were included in the protocol: (1) opening, (2) introductory, (3) key, (4) probe, and (5) ending. An *opening question* allows the participants to acclimate to the group setting and is only loosely connected to the research topic. This was followed by an *introductory question,* which broadly presents the research topic, allowing the participants to offer initial and overarching attitudes before focused questioning. A *key question* is a focused inquiry specifically worded to answer a project goal or research question. I designed three key questions to match the three research questions. While a single key question is usually offered by the moderator to start discussion, *probe* questions are used to further this discussion, asking for clarification of ideas and experiences offered by participants (Monmonier and Gluck 1994). Each focus group session concluded with a broad *ending question*, allowing participants to offer closing comments. The complete focus group protocol for the two sessions, detailing the five categories of questioning used in the study, is available in Table 1.

The second step of the moderating stage is to choose and train the moderator. The *moderator*, or facilitator, is responsible for keeping the participants focused around the key questions while allowing open-ended discussion. It is essential for the moderator to have comprehensive expertise concerning the discussion topics (Harrower et al. 2000). Cameron (2005, 124) states that, "In academic research it is often the researcher, who is familiar with the aim of the research and the purpose of the focus groups, who is best positioned to fill this role." However, Krueger (1998a) stresses the importance of a neutral moderator, arguing that allowing the principle investigator to act as the moderator may bias participant response. Such investigator bias is a serious concern when the goal of a focus group is to test a software application built by the moderator, as participants may be reserved in offering negative feedback when the lead developer of a product is present (Kessler 2000).

I acted as the moderator for both focus group sessions. This decision was justified by four points. First, proper moderation of the focus groups required expertise in both geographic information uncertainty and floodplain mapping; my preparation for the research project placed me in the unique position of having ample understanding of both topics. Second, while I could have secured the services of an expert in either geographic information uncertainty or floodplain mapping, I would have had the task of training the moderator in his or her deficient area. Given the time and resource constraints of the research, it was assumed that a trained moderator would still have major gaps in one of the two areas despite best efforts. Third, there is ample precedence in the GIScience literature of allowing the principle investigator to act as moderator (e.g., Monmonier and Gluck 1994; Harrower et al. 2000; Kessler 2000). Finally, I believed that investigator bias was less of a concern for this research because I was not testing an original concept or product; I was instead determining the appropriateness of a typology offered by other scholars and therefore had no explicit opinion about its legitimacy.

3.4. Analysis

The first step of the analysis is transcription of the focus group discussions. The audio recordings were transcribed by a student in the Psychology Department of the University of Wisconsin-Madison with training and experience in transcribing focus groups. The transcriber was instructed to place each sentence from the discussion into a separate row in a Microsoft Excel database and then to mark in a second column the participant contributing the comment using an arbitrary ID. The margin coding approach was used to analyze the transcriptions. In *margin coding*, a set of important themes are identified by the researcher and given number or letter codes (Bertrand et al. 1992). The person conducting the coding analyzes each unit in the transcription and applies a particular code in the margin next to the statement if the theme is present. To reduce bias, it is best if a person other than the moderator performs the coding so that the coder treats the entirety of the transcription equally, rather than focusing upon selected points of interest. The same student that transcribed the focus group sessions also completed the coding, but was not given the list of themes until after the transcriptions were completed and reviewed. Following Cameron (2005), a unique coding scheme based on the MacEachren et al. (2005) typology was developed for each of the three key questions (Table 2), producing three separate coded transcriptions.

4. Reporting of Results and Discussion

The coding schemes described in the previous section were used to organize statements for the synoptic style of reporting found in Monmonier and Gluck (1994). This kind of reporting provides a descriptive summary for each code and uses direct quotes that exemplify important issues or opinions. In addition to synoptic reporting, I employed Krueger's (1998b) frequency and extensiveness for the second research question. *Frequency* is a reporting of the number of times an individual code is present in a transcript and *extensiveness* is a reporting of the number of participants that mention an individual code, regardless of frequency. Using frequency and extensiveness, it is possible to compare the nine categories of uncertainty using simple metrics as a supplement to the more detailed synoptic style of reporting. Results for the three research questions are reported in the following three subsections.

4.1 The involvement of uncertainty in floodplain mapping

The purpose of the first key question was to provide an understanding of how geographic information uncertainty is involved in floodplain mapping. Comments from the focus groups concerning floodplain mapping practice naturally separated into three broad groupings: (1) those relating to the creation of floodplain information (including all subsequent transformations), (2) those relating to the representation of floodplain information, and (3) those relating to the use of floodplain information in support of decision-making. These three groupings described general steps for arriving at a decision outcome about reality using geographic information, as displayed in the Figure 2 flowchart. This flowchart parallels closely to the Agumya and Hunter (2002) decision-making framework. However, I propose a revision to the Agumya and Hunter (2002) framework to include the intermediate step of 'Uncertainty Representation' between 'Data Uncertainty' and 'Decision Uncertainty' to match the prototypical process described by Figure 2.

The proposed revision to the Agumya and Hunter (2002) framework in relation to the Figure 2 flowchart is provided in Figure 3.

This flowchart was then modified to track the involvement of uncertainty in each of the three steps, allowing for a reconstruction of the way in which uncertainty is utilized in the domain of floodplain mapping (Figure 4a). There are four possible pathways through the flowchart, differing in the step at which uncertainty is no longer involved and assuming that once geographic information uncertainty is not involved, measures of the uncertainty taken during creation can no longer be reliably reincorporated into following steps (Figures 4b through 4e). It is important to note that the flowcharts in Figure 4 are highly simplified and purposefully gloss over many aspects of the geographic information lifecycle and the decision-making process.

Participants in each session displayed a clear understanding of the uncertainty in the information created by their organizations. Multiple participants noted that much of the available floodplain information was digitized from non-orthorectified paper maps. The uncertainty in this process was acknowledged, with one participant stating that the digitized line is only “an approximation,” another saying that “We created a line [by digitizing a firm’s maps] and of course it didn’t always fit right,” and another adding that this process of creation was a “worst case scenario” for introducing uncertainty. However, participants reported a greater confidence in natively digital geographic information compiled more recently, stating that the new commissions are forced to conform to “a list of criteria” or threshold of acceptable uncertainty provided by FEMA. However, there was still an understanding among the participants that uncertainty is present in these newer datasets, with one participant stating “Looking at the products that we have from FEMA right now, you're always uncertain” and another stating that there are “places [in the FEMA floodplain information] where they don’t even overlay a stream.”

Despite acknowledging the involvement of uncertainty in the creation step, there was a serious resistance reported by participants against representing uncertainty visually. Participants could not recall a floodplain map they have encountered that indicated uncertainty beyond a textual disclaimer. One participant argued that “You just have to assume that the line you draw [on the map] is a hard and fast line …you’ve got to put a line somewhere.” Figure 4c portrays this situation in which uncertainty is deemed important to consider during geographic information creation, but not for representation.

A first explanation of this resistance is that the mapmaker is hesitant to represent or discuss the uncertainty of the geographic information for fear of undermining it. Participants indicated that once geographic information meets the FEMA uncertainty criteria, the representation can be used as if it was fully certain. One participant said that “If you put [uncertainty] on the map it would probably draw undue attention.” A second participant stated strongly that representing uncertainty “would bring to the forefront the questionability of the map…that's always an issue anyways, but now you're putting a logo on it and saying ‘Don't forget to ask me about this.’” These comments follow Mowrer’s (1999, 3) remark that “Perhaps the worst nightmare of a natural resources manager is to appear ‘uncertain’ to the public, or to admit that there is ‘error’ in the decision process being presented.”

A second explanation is that domain experts are not convinced that end users are capable of understanding the uncertainty depictions. The perception of incapability is predicated on two interlinked points: (1) assumed inability of users to understand the concept of uncertainty and (2) assumed inability of users to understand a representation of uncertainty. The participants worried that decision-makers would conflate the idea that “uncertainty means use with caution” with “uncertainty equals worthless,” leading the decision-makers to “throw out” the information completely. Further, participants did not believe that current methods of representation effectively communicate the presence or magnitude of uncertainty to decision-makers, particularly those that are novice map readers. However, one participant did offer that he "would be all for anything about uncertainty on the map as long as it was really understandable for the layman."

A final explanation is revealed by the way in which floodplain maps are used for decision-making support. There was a general assumption that decision-makers did not want to know the degree of uncertainty in the geographic information. One participant offered the assumption that if "You paid enough for it to be right…Of course it's right." A second participant, speaking as a decision-makers, stated that "maybe sometimes we don't want to know that information." Finally, a third participant described the desire to reduce a decision to a binary task, stating that "if your decision has to be an absolute yes/no, putting the uncertainty on the map puts you in the position [where] you're the one that has to make the call…it has added problems to your life." These comments suggest that decision-makers purposefully treat 'best available' information as 'best possible' and try to reduce a decision to a simple yes/no, closely reflecting the Agumya and Hunter (2002) threshold comparison of estimated and acceptable risk.

The Agumya and Hunter (2002) decision-making framework and participant comments suggest that the pathway exhibited in Figure 4c is less than ideal. By removing representations of uncertainty from the floodplain map, the presence and magnitude of uncertainty in the geographic information cannot be communicated to decision-makers. The decision-makers may then approach the decision as if there was full certainty, incorrectly placing undo confidence in their risk response. As one participant noted, "the fact that we have a line on a map, now people are viewing that as gospel." There are dangers in incorrectly assuming conditions of certainty during a decision, with one participant offering that “people who believe that line is absolute and don't see the uncertainty are making absolute regulations based on it... if people understood that the line wavered, maybe those regulations would be tempered and not so absolute.” It is thus necessary to transition from an involvement of uncertainty illustrated in Figure 4c to one illustrated in Figure 4e to improve the decision-making in the domain of floodplain mapping.

4.2. The appropriateness of the MacEachren et al. (2005) typology for floodplain mapping

The purpose of the second key question was to determine the appropriateness of the MacEachren et al. (2005) typology when applied to the domain of floodplain mapping. Establishing the appropriateness of the MacEachren et al. (2005) typology was an important first step prior to investigating the relative influence of each category on the decision-making process. Because the literature reports a tendency for participants to agree with materials presented by the moderator (e.g., Kessler 2000; Cameron 2005), the first key question also served the purpose of brainstorming uncertainty categories in floodplain mapping before prompting of the MacEachren et al. (2005) typology. Table 3 provides a summary of the unprompted and prompted frequency and extensiveness of each category.

Six categories were positively identified as components of uncertainty without prompting of the MacEachren et al. (2005) typology. After prompting, all nine categories were described numerous times as a component of geographic information uncertainty in the domain of floodplain mapping. A minimum frequency of ‘14’ is surprisingly high considering the small participant number. Unprompted extensiveness showed less support for the MacEachren et al. (2005) typology, with only accuracy/error and currency individually mentioned by the majority of participants. However, after circulation of the MacEachren et al. (2005) typology, all nine categories exhibited extensiveness values of either ‘5’ or ‘6’. This strongly supports the appropriateness of the MacEachren et al. (2005) typology for application to the domain of floodplain mapping, as each category was considered a component by either a large majority or all participants. The large frequencies, both unprompted and prompted, and the near complete prompted extensiveness values suggest that each of the nine categories are legitimate components of uncertainty in the domain of floodplain mapping.

Further support is provided by comments about the MacEachren et al. (2005) typology after circulation. When presented with the typology, one participant stated "well I think they all play a role to some degree." Another participant added that "the rest of these things on the list [not yet discussed before circulation of the typology] you could pretty much make the case for." Drawing from an experience in creating floodplain information, a third participant stated "I'm sure at the time that ours was compiled all nine of these were taken into consideration…and it probably wasn't approved either until all nine of these were taken into consideration, both at the FEMA, state, and local level." Finally, a fourth participant stated that "I think that it's such an important thing to get right that if you don't have the time to take all nine of these things into consideration, then you just have to ask to have the deadline extended to get it right."

During this portion of discussion, I was particularly interested in the identification of categories included in the typology that do not match the reality of floodplain mapping or categories not currently included in the typology. A surprising finding of the study was that there was not a single comment in either session arguing against the inclusion of an uncertainty category listed in the MacEachren et al. (2005) typology. However, both sessions did reference a tenth possible uncertainty category when discussing the importance of the requester on the certainty of the geographic information, particularly in the case of the FEMA certainty criteria. The participants argued that the same firm could release two datasets mapping the same floodplain, but that the datasets could have a varying degree of certainty based on the quality needs of the client. In this example, the credibility of the geographic information was equally dependent upon the information producer and the information requester. Such a situation is commonly considered a component of information quality or fitness-of-use (Longley et al. 2005). Rather than adding a new category of uncertainty for the needs of the information requester, I believe that the definition of credibility should be amended to include both the information source and requester.

4.3. Determining the relative influence of different uncertainty categories

The purpose of the final key question was to determine the relative influence of each category of uncertainty on the decision-making process. Focus group participants unanimously agreed that accuracy/error was the most influential category of uncertainty on decision-making. One participant stated “The number one thing that people are bringing up...is the accuracy issue, and, hands down, that is the number one issue.” Similarly, a second participant said that “accuracy is the one thing that [decision-makers] always come for.” A third participant added the notion that perhaps accuracy is the only category of uncertainty worth representing, providing an example where policy makers may no longer make absolute regulations when the floodplain boundary is represented as a gradient rather than a single line.

The categories precision and currency also were identified as highly influential. While the concept of precision was not mentioned prior to circulation of the MacEachren et al. (2005) typology, participants recognized this category as influential in both sessions following prompting. One participant in the first session stated that “accuracy, I would also put precision under the same which is second here, are both cases that, those are the things that the engineers and the people are really looking under the hood for.” This participant went on to say that

"Accuracy and precision is in the specs that you did... So you're absolutely right, that's what they go after in any mapping, floodplain or anything else." Finally, all three participants in the second session agreed that numbers "one and two" in the typology (i.e., accuracy/error and precision) were the most influential to decision-making.

Participants also reported a strong influence of currency on decision-making, with several comments indicating that currency was near to or equaled the influence of accuracy/error. One participant began discussion on influence stating that "I think you start with currency." A different participant stated that "accuracy is one thing, but currency, and that is obviously on your list [of categories], is a big deal." This participant goes on to say that "So my take on that is that accuracy is number one but certainly currency would be two." There were no statements describing any of the remaining six categories of uncertainty from the MacEachren et al. (2005) typology as non-influential; I instead deemed the remaining six categories as secondary in influential due to the lack of response.

5. Conclusions

The focus groups provided initial evidence that there is a strong disconnect between the actual practice of geographic information uncertainty representation (illustrated by Figure 4c) and the ideal practice recommended by the literature (illustrated by Figure 4e). Rather than representing the uncertainty measurements recorded during information creation for reference during decision-making, uncertainty recordings are used solely as a quality check during creation of the floodplain information. This usage of uncertainty information is analogous to the original purpose of FDA drug ratings in the 1970s. FDA drug ratings were initially intended for use in determining the safety of a drug before allowing its sale by pharmaceutical companies. Drugs that met established standards of safety, although still containing health risks, were passed by the

FDA for sale and did not require a detailed explanation of the risks uncovered by the FDA while researching the drug. It was found that the lack of explicit risk descriptions caused consumers to inappropriately judge a drug's level of risk as a safe/unsafe binary decision. Current laws now require detailed descriptions of the health risks of a drug to be printed on the bottle (i.e., an uncertainty representation), supporting a fully informed decision of its safety by the consumer.

The focus groups also produced support for the MacEachren et al. (2005) typology as an appropriate model of uncertainty categorization in the domain of floodplain mapping. Results did not suggest the need to add a category to the typology or delete a category from it, but did reveal a necessary revision of the definition of credibility to include the needs of the information requester alongside those of the information provider. Due to several limitations of the study, it is only possible to suggest (rather than prove) that the MacEachren et al. (2005) typology is a valid listing of geographic information uncertainties. First, it is important to reiterate that the nature of focus group research only allows for claims concerning the specific case study domain. As the typology is meant to be a general model view of uncertainty in any geographic information, it would be interesting to see how the results from the domain of floodplain mapping extend to other domains and groups of decision-makers. Second, only the MacEachren et al. (2005) typology was circulated in the second half of the focus groups. Circulation of multiple typologies with discussion of their relative advantages and limitations would have provided more insight into the validity of the MacEachren et al. (2005) typology itself, although at the expense of collecting information on the other research questions. For these reasons, future investigation is necessary before declaring the MacEachren et al. (2005) typology, as it is currently defined, as the final word on uncertainty categorization.

Finally, focus group participants identified accuracy/error, precision, and currency as the most influential in the decision-making process. Understanding which categories of uncertainty are necessary for informed decision-making is a critical first step towards improving the representation of uncertainty for support of decision-making. In many mapping contexts, particularly static or print mapping that cannot take advantage of interactivity, it is plausible to only represent a single uncertainty category atop the geographic information. The results of the focus groups suggest that under such design limitations, accuracy/error, precision, and currency are the most important to communicate to the decision-makers.

As uncertainty is “not just a flaw that needs to be excised” (Couclelis 2003, 166), it is essential that researchers continue to adapt traditional methods of representation to uncertainty and to develop novel ways of representation specific to uncertainty. Following Agumya and Hunter (2002), acknowledgement of the uncertainty intrinsic to the input dataset is necessary for making a fully informed decision. However, because the decision-making process often is compartmentalized, with one group of individuals creating the geographic information, a second producing the representations, and a third making the decision, it is possible that the end decision-makers is unaware of the limitations of the underlying geographic information if uncertainty representations are not included. The presence of numerous uncertainty categories only complicates the problem of proper representation further. Yet, if support for informed decision-making is the end goal, it is a problem that necessitates a well researched solution.

6. Acknowledgements

I would like to thank Mark Harrower for providing input on the focus group design and Alan MacEachren for providing feedback on the manuscript after the research was completed. I wish to thank Meena Pandian for her assistance with transcribing and codifying the focus group audio recordings. I would also like to thank Ann Barrett and the Wisconsin Land Information Association for circulating a call for participation in the research to conference attendees and for providing time and space for the focus groups during the 2007 WLIA conference. Finally, I wish to thank the anonymous reviewers for their time and insightful input.

Bibliography

Agumya, A., and G. Hunter. 2002. Responding to the consequences of uncertainty in geographical data. *International Journal of Geographical Information Science* 16(5): 405-417.

Beard, K., and W. Mackaness. 1993. Visual access to data quality in geographic information systems. *Cartographica* 30(2-3): 37-45.

Bertrand, J.T., J.E. Brown, and V.M. Ward. 1992. Techniques for analyzing focus group data. *Evaluation Review* 16(2): 198-209.

Brown, J.B. 1999. The use of focus groups in clinical research. In: B.F. Crabtree and W.L. Miller (eds) *Doing qualitative research*. Thousand Oaks, CA: Sage Publications.

Buttenfield, B. 1993. Representing data quality. *Cartographica* 30(2-3): 1-7.

Cameron, J. 2005. Focussing on the focus group. In: I. Hay (ed) *Qualitative research methods in human geography*. Melbourne, Australia: Oxford University Press.

Carroll, L. 1982. Sylvie and Bruno Concluded. In: E. Guillaro (ed) *The complete illustrated works of Lewis Carroll*. New York, NY: Avenel Books. Originally published in 1893.

Couclelis, H. 2003. The certainty of uncertainty: GIS and the limits of geographic knowledge. *Transactions in GIS* 7(2): 165-175.

Edwards, L.D., and E.S. Nelson. 2001. Visualizing data certainty: A case study using graduated circle maps. *Cartographic Perspectives,* 38(Winter): 19-26.

Harrower, M., A.M. MacEachren, and A. Griffin. 2000. Developing a geographic visualization tool to support earth science learning. *Cartography and Geographic Information Science* 27(4): 279-293.

Harrower, M. 2003. Representing uncertainty: Does it help people make better decisions? In: *Proceedings of the UCGIS Geospatial Visualization and Knowledge Discovery Workshop.* November 18-20, 2003: Landsdowne, VA. CD-ROM.

Kessler, F. 2000. Focus Groups as a means of qualitatively assessing the U-boat narrative. *Cartographica* 37(4): 33-60.

Krueger, R. 1998a. *Focus group kit #3: Developing questions for focus groups.* Thousand Oaks, CA: Sage Publications.

Krueger, R. 1998b. *Focus group kit #6: Analyzing and reporting focus group results.* Thousand Oaks, CA: Sage Publications.

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

Longley, P.A., M.F. Goodchild, D.J. Maguire, and D.W. Rhind. 2005. *Geographic information systems and science, 2nd edition*. Hoboken, NJ: John Wiley & Sons.

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

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): 139-160.

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

Monmonier, M., and M. Gluck. 1994. Focus groups for design improvement in dynamic cartography. *Cartography and Geographic Information Systems* 21(1): 37-47.

Morgan, D.L. 1998. *Focus group kit #2: Planning focus groups*. Thousand Oaks, CA: Sage Publications.

Mowrer, H.T. 1999. Accuracy (re)assurance: Selling uncertainty assessment to the uncertain. In: K. Lowell and A. Jaton (eds) *Spatial accuracy assessment: Land information uncertainty in natural resources*. Chelsea, MI: Ann Arbor Press.

Patton, M.Q. 1990. *Qualitative evaluation and research methods*. Newbury Park, CA: Sage Publications.

Roth, R.E. 2009. The impact of user expertise on geographic risk assessment under uncertain conditions. *Cartography and Geographic Information Science*

Sinton, D. 1978. The inherent structure of information as a constraint to analysis: Mapped thematic data as a case study. In: D. Dutton (ed.) *Harvard papers in GIS #7.* Cambridge, MA: Harvard University.

Thomson, J., B. Hetzler, A. MacEachren, M. Gahegan, and M. Pavel. 2005. Typology for visualizing uncertainty. In: *Proceedings of the IS&T/SPIE Symposium on Electronic Imaging, Conference on Visualization and Data Analysis.* January 16-20, 2005: San Jose, CA. CD-ROM.

Twohig, P.L., and W. Putnam. Group interviews in primary care research: Advancing the state of the art or ritualized research? *Family Practice* 19(3): 278-284.

Valentine, G. 1997. Tell me about...: using interviews as a research methodology. In: R. Flowerdew and M. Martin (eds) *Methods in human geography: a guide for students doing research projects*. Essex, England: Addison Wesley Longman.

Zhu, A.X. 2005. Research issues on uncertainty in geographic data and GIS-based analysis. In: R. McMaster and L. Ursery (eds) *A research agenda for geographic information science*. New York, NY: CRC Press.

Figures:

Figure 1: The decision-making process under uncertain conditions, printed in Agumya and Hunter (2002, 407). Because the input dataset contains uncertainty, uncertainty is present throughout the decision-making process.

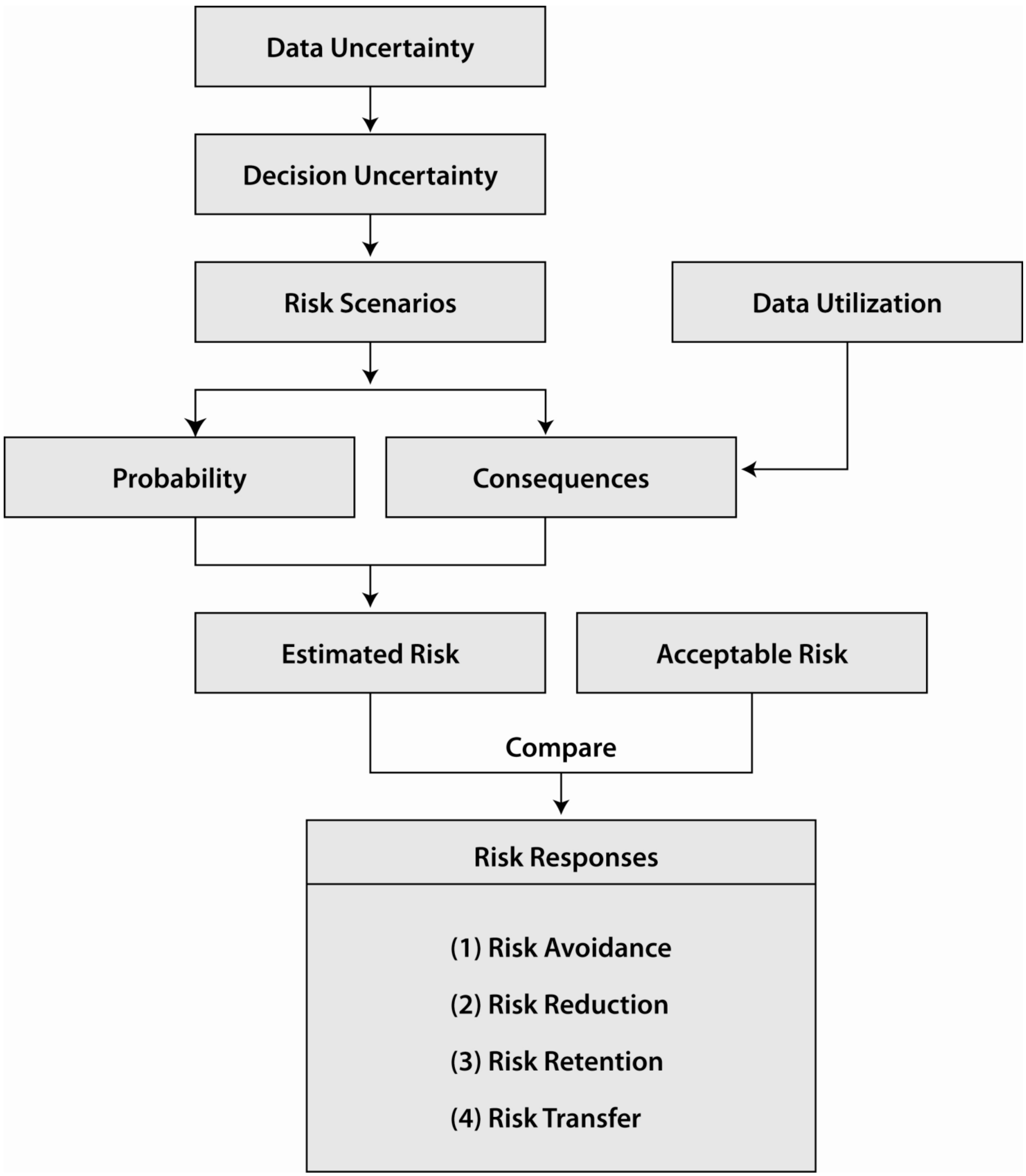


Figure 2: A flowchart describing the three major steps necessary for arriving at a decision outcome about reality using geospatial data. Statements concerning the first key question fell naturally into three groupings: (1) those concerning geographic information creation (including all transformations), (2) those concerning the representation of the geographic information, and (3) those concerning the use of the representation in support of decision-making.

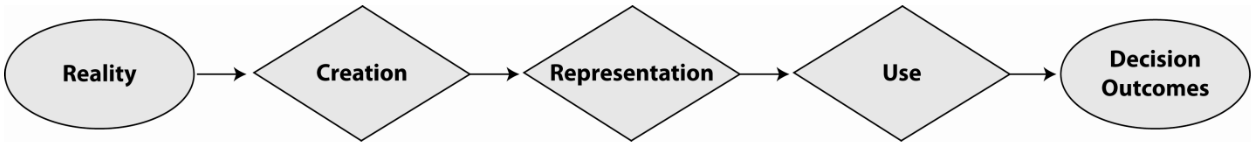


Figure 3: The Figure 2 flowchart in relation to the Agumya and Hunter (2002) decision-making framework. It is important to note that the framework was adjusted to include an intermediate ‘Uncertainty Representation’ step between ‘Data Certainty’ and ‘Decision Certainty.’

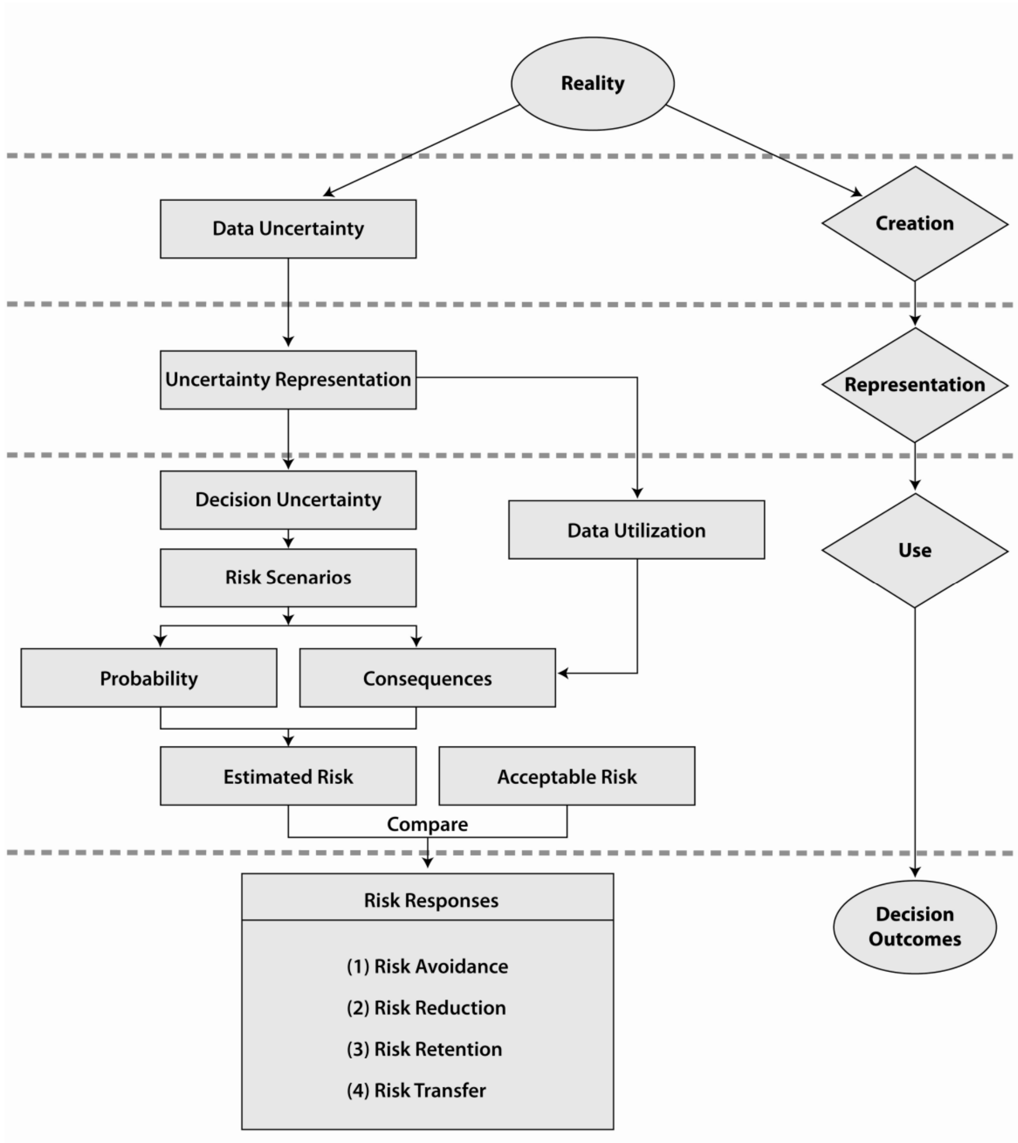
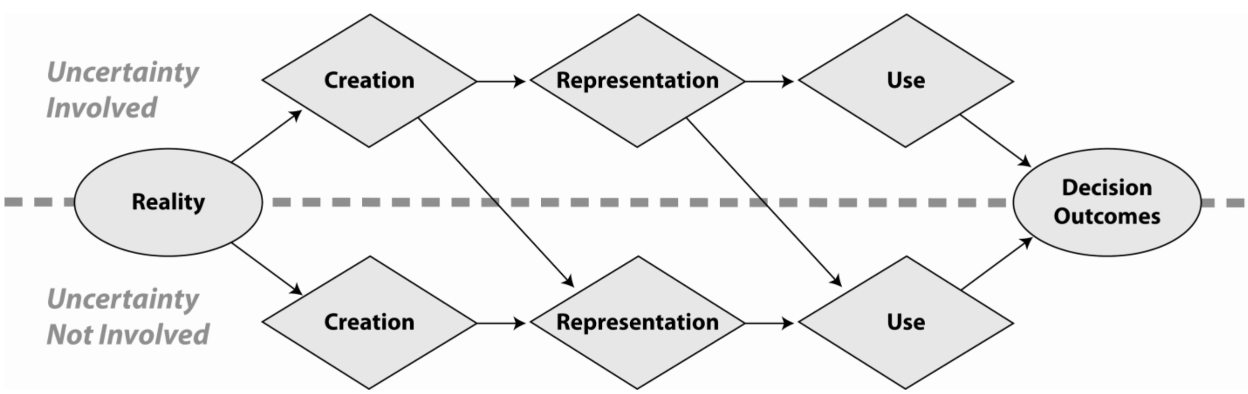


Figure 4a: The flowchart describing the three major steps necessary for arriving at a decision outcome about reality using geographic information, revised to acknowledge the possible involvement of uncertainty at each of the three steps.



Figures 4b-4e: The four possible pathways through the flowchart. Although the literature states that the ideal pathway is similar to Figure 4e, the focus groups revealed a practice similar to Figure 4c.

Figure 4b:

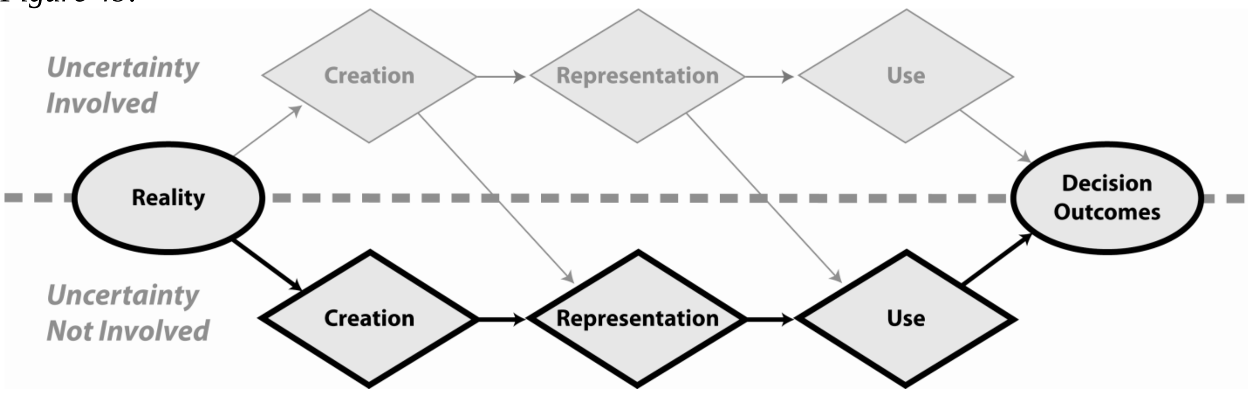


Figure 4c:

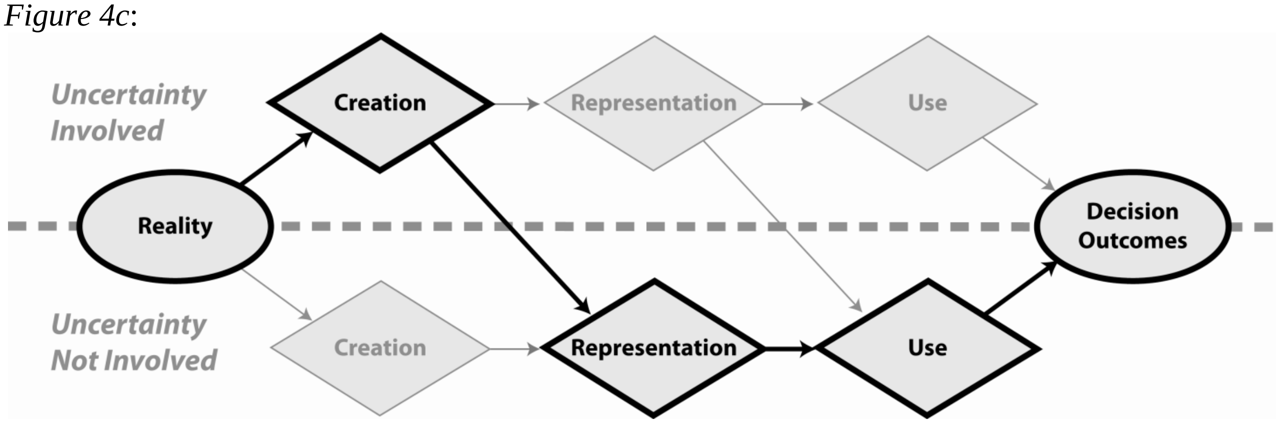


Figure 4d:

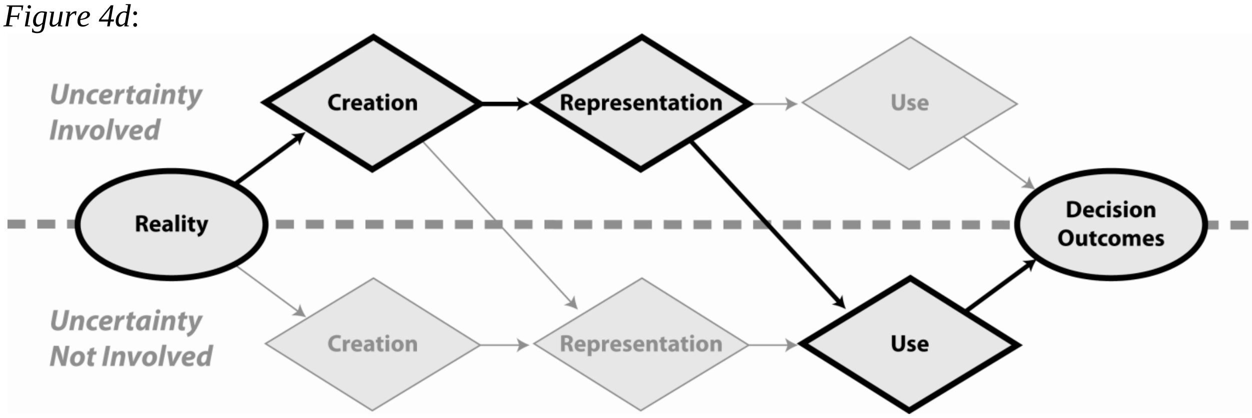
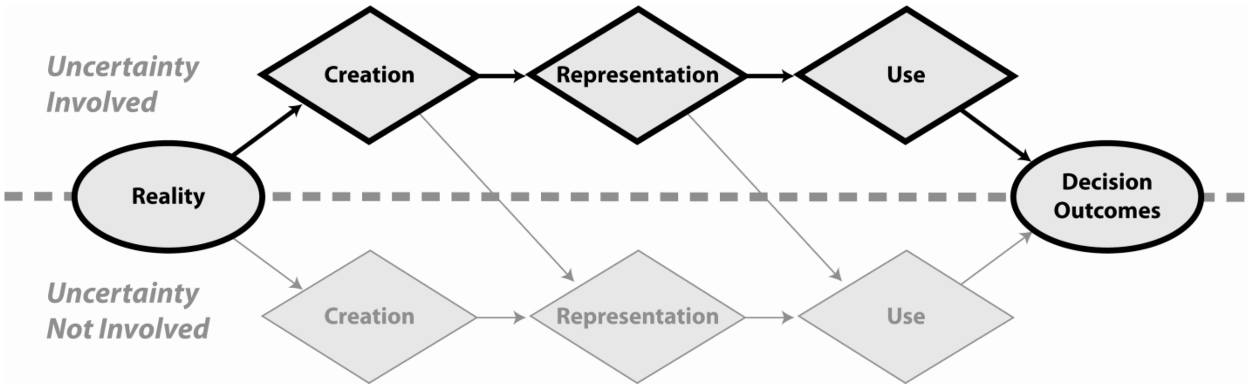


Figure 4e



Tables

*Table 1:*The focus group protocol used for both focus group sessions. Five types of questions were included: (1) opening, (2) introductory, (3) key, (4) probe, and (5) ending.

Question Type	*Question Description*
Opening	Please introduce yourself to the group and briefly describe your work responsibilities with regards to floodplain mapping.
Introductory	Describe the concept of uncertainty.
Key #1	How has uncertainty entered into floodplain mapping in your experience? *Probe: Can you think of any other ways in which uncertainty could enter into floodplain mapping?* *Probe: Has uncertainty affected the way you collect floodplain information or make floodplain maps?* *Probe: How do you represent uncertainty in your floodplain maps?*
Key #2	Have you encountered any of these types of uncertainty in your work experience? *Probe: Do any of these types not apply to the uncertainty that is found in floodplain mapping?* *Probe: Are there types of uncertainty in floodplain mapping not included in this list?* *Probe: Does knowing about this list change the way you would work with uncertainty in floodplain mapping?*
Key #3	When making a decision using a floodplain map, which uncertainty type(s) do you weight most heavily? *Probe: When making a decision using a floodplain map, which uncertainty type(s) do you think is/are the most important to represent?* *Probe: Is this type(s) always the most influential to a decision, or is influence situational?* *Probe: If you were informed that a high degree of uncertainty exists in a particular type(s), which type(s) would cause you to approach a decision with the most caution?*
Ending	Based upon this discussion, do you have any closing comments or questions for me to clarify?

Table 2: The coding scheme applied to the focus group transcriptions for each of the three key questions.

Table #2a: Coding for Question #1

Code	*Description*
C-no	statements describing the creation of floodplain information that do not involve uncertainty (irrespective of category)
C-yes	statements describing the creation of floodplain information that involve uncertainty (irrespective of category)
R-no	statements describing the representation of floodplain information that do not involve uncertainty (irrespective of category)
R-yes	statements describing the representation of floodplain information that involve uncertainty (irrespective of category)
U-no	statements describing the use of representations of floodplain information that do not involve uncertainty (irrespective of category)
U-yes	statements describing the use of representations of floodplain information that involve uncertainty (irrespective of category)

Table #2b: Coding for Question #2

Code	Description
acc/err	statements that positively identify accuracy/error as a category of uncertainty
prec/res	statements that positively identify precision as a category of uncertainty
comp	statements that positively identify completeness as a category of uncertainty
cons	statements that positively identify consistency as a category of uncertainty
line	statements that positively identify lineage as a category of uncertainty
curr	statements that positively identify currency as a category of uncertainty
cred	statements that positively identify credibility as a category of uncertainty
subj	statements that positively identify subjectivity as a category of uncertainty
inter	statements that positively identify interrelatedness as a category of uncertainty
xxx	statements that positively identify a category not identified by the MacEachren et al. (2005) typology

Table #2c: Coding for Question #3

Code	Description
acc/err++	statements reflecting a high degree of influence on decision-making for the uncertainty category accuracy/error
acc/err--	statements reflecting a low degree of influence on decision-making for the uncertainty category accuracy/error
prec/res++	statements reflecting a high degree of influence on decision-making for the uncertainty category precision/resolution
prec/res--	statements reflecting a low degree of influence on decision-making for the uncertainty category precision/resolution
comp++	statements reflecting a high degree of influence on decision-making for the uncertainty category completeness
comp--	statements reflecting a low degree of influence on decision-making for the uncertainty category completeness
cons++	statements reflecting a high degree of influence on decision-making for the uncertainty category consistency
cons--	statements reflecting a low degree of influence on decision-making for the uncertainty category consistency
line++	statements reflecting a high degree of influence on decision-making for the uncertainty category lineage
line--	statements reflecting a low degree of influence on decision-making for the uncertainty category lineage
curr++	statements reflecting a high degree of influence on decision-making for the uncertainty category currency
curr--	statements reflecting a low degree of influence on decision-making for the uncertainty category currency
cred++	statements reflecting a high degree of influence on decision-making for the uncertainty category credibility
cred--	statements reflecting a low degree of influence on decision-making for the uncertainty category credibility
subj++	statements reflecting a high degree of influence on decision-making for the uncertainty category subjectivity
subj--	statements reflecting a low degree of influence on decision-making for the uncertainty category subjectivity
inter++	statements reflecting a high degree of influence on decision-making for the uncertainty category interrelatedness
inter--	statements reflecting a low degree of influence on decision-making for the uncertainty category interrelatedness

Table 3: A summary of the frequency and extensiveness of statements in support of the appropriateness of the MacEachren et al. (2005) typology.

Category	Frequency			Extensiveness		
	Unprompted	Prompted	Total	Unprompted	Prompted	Total
Accuracy/Error	21	68	89	5	6	6
Precision/Resolution	0	32	32	0	6	6
Completeness	0	15	15	0	5	5
Consistency	0	21	21	0	5	5
Lineage	2	16	18	2	5	5
Currency	8	40	48	4	6	6
Credibility	2	26	28	2	5	5
Subjectivity	7	32	39	3	6	6
Interrelatedness	1	13	14	1	5	5