“I don’t know what’s going on there”: The Use of Spatial Transformations to Deal With and Resolve Uncertainty in Complex Visualizations

Susan Bell Trickett, George Mason University

J. Gregory Trafton, Naval Research Laboratory

Lelyn Saner, University of Pittsburgh

Christian D. Schunn, University of Pittsburgh

Introduction

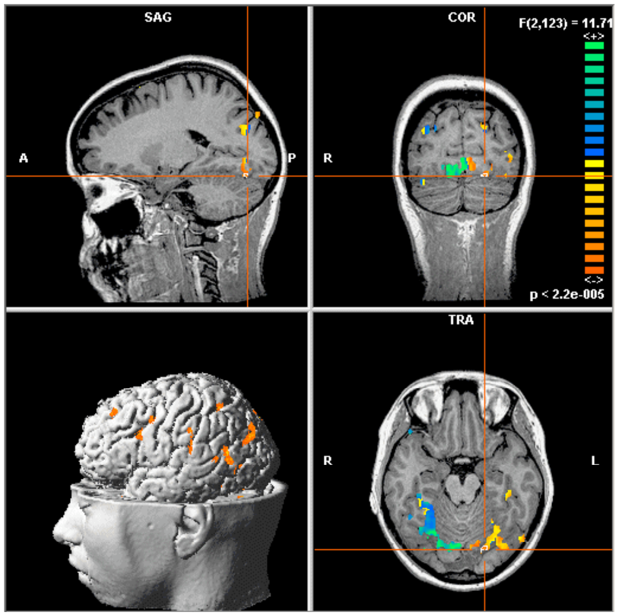
Imagine a meteorologist preparing a weather forecast. In addition to years of experience and a vast store of domain knowledge, the forecaster has access to satellite images, to computer–generated weather models and programs to display them in a variety of ways, and to an assortment of special-purpose tools that provide additional task-relevant data. There is no shortage of data, yet despite this array of resources, the task remains very challenging. One source of complexity is the uncertainty inherent in these data, uncertainty that takes many forms. Why are two weather models making different predictions? Are the models based on many observations or just a few? Are there enough observations in a given model to trust it? Is one model more reliable than another in certain circumstances, and if so, what are they? Which one, if either, should be believed? How long ago were these data collected? How have things changed since the data were originally displayed? What is the real location of this front, and how is it affected by other changing variables, such as wind direction and speed, which may also have changed?

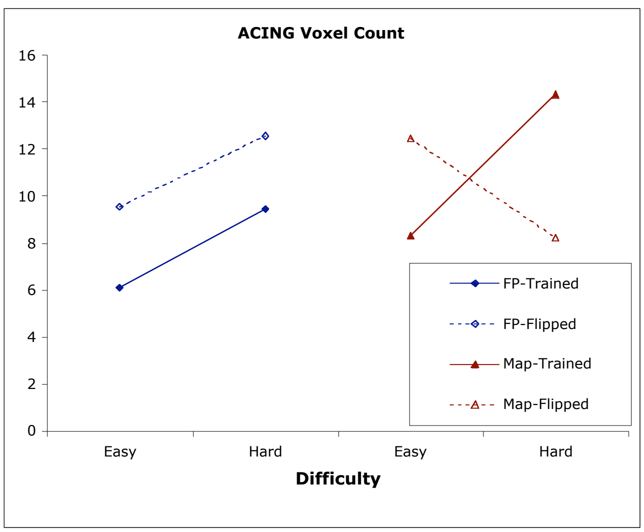
To complicate matters further, the uncertainty in the data is not explicitly represented; rather, the visualizations indicate that the data *are* exactly as they *appear*. The visualizations thus invite the forecaster to map uncertain data to certain values, yet to do so would most likely lead to erroneous predictions. How does he or she manage this incongruity, in order to develop the most accurate forecast possible?

This example illustrates the basic question we investigate in this paper: how do people, especially experts, deal with uncertainty in highly spatial domains, when the data are inherently uncertain but the tools actually *display* very little uncertainty? Experts in many domains depend on complex visualizations that use spatial representations of data, in areas as diverse as military operations (e.g., testing and evaluation of electronic warfare systems and of techniques to counter antiship missile threats, mission rehearsal prior to combat), geosciences (e.g., weather analysis and forecasting, geology, environmental science, oceanography) and scientific visualization (e.g., neuroscience, computational fluid dynamics, molecular biology, medical research and practice). In each of these examples—and there are many more we could cite—the practitioner must contend with uncertainty in the data. We first examine how uncertainty affects operations in three representative domains, submarine operations (military), meteorology (geoscience), and fMRI research (scientific visualization), in which dealing with uncertainty is a critical component of the task. We then investigate how experts in two of these domains, meteorology and fMRI, manage uncertainty as they perform problem-solving activities and make decisions as part of their regular task performance.

The sources of uncertainty in these three domains are many and varied. Uncertainty is inherent in the submarine world due to the nature of the primary sensory system, passive sonar. The causes of uncertainty are the ocean environment (the physics of sound transmission through ocean water, interactions with the bottom, wave action, noise caused by ocean creatures and ships, etc.), the under-determined nature of target motion analysis (TMA) from bearings-only measurements, and the unpredictability of human actions and intentions (both unintentional and intentional deception). These problems have become even more critical with the emphasis on operations in the crowded and noisy littoral regions, and uncertainty is the primary cause of error and delayed action. In meteorology, as outlined above, there is considerable uncertainty in Meteorology and Oceanography (METOC) data, both in observations and models. Weather models are based on underlying assumptions that may or may not be accurate, and they may depend on unreliable or sparsely sampled observations to make their predictions; satellite images do not portray current conditions, but rather show "truth" as it was some time in the recent past; weather-related variables are depicted as having absolute values or locations, whereas in reality, only approximations can be displayed. Likewise, in fMRI, several factors contribute to an overall high level of uncertainty. Irrelevant areas of the brain may be activated by subjects' off-task thoughts. The spatial resolution of the display itself is considerably coarser than neurons or even assemblies of neurons. The measurements themselves can be systematically biased by a variety of factors (e.g., the closeness to the skull, deformations in areas near the nasal passages due to breathing). In addition, the processes of neurons themselves are thought to be stochastic and the neuronal processes happen at a faster pace than the temporal resolution of fMRI. Moreover, to deal with the measurement noise, the analysis of the data typically averages data across several seconds of time or across many trials.

Despite the many sources of uncertainty in these domains, the visualizations often do not explicitly display the uncertainty, but may rather present data in a much less ambiguous fashion than is congruent with reality. For example, Figure 1 shows examples of three typical different visualizations of fMRI data. Moving clockwise, from the top, they show the degree of activation, indicated with a color scale superimposed over a gray-scale structural brain image in three different planar slices and a surface cortex map; a graph of the number of activated voxels in an area as a function of various condition manipulations; and a table of the number of activated voxels in different brain areas (Regions of Interest) as a function of different conditions. Note the lack of uncertainty represented in the display. For example, in the first visualization, the lit areas signifying neural activity are clearly bounded, and even the different colors are unambiguously discrete. Thus the visualization suggests a precision in the mapping between location in the brain and level of neural activity that, in fact, is unlikely to reflect the actual activity within the brain.





ROI	TrainedEasy Count	TrainedHar Count	FlippedEasy Count	FlippedHard Count
1	2	3	4	5
ACING	8.33	14.33	12.44	8.22
CALC	72.33	87.56	75.89	103.22
LCBELL	21.11	33	27.22	28.89
RCBELL	19.44	38.44	28.33	33.67
LCBELL+RCBELL	40.56	71.44	55.56	62.56
LDLPFC	**32.78**	**51**	**38.22**	**40.56**
RDLPFC	**37.56**	**61.67**	**43.44**	**57.33**
DLPFC	**70.33**	**112.67**	**81.67**	**97.89**
LFEF	5.33	8.22	4.67	7.33
RFEF	3.44	5.56	3.78	6.56
LFEF+RFEF	8.78	13.78	8.44	13.89
LHG	12.67	1.44	5.56	1.56
RHG	14	3	6	3.78
LHG+RHG	26.67	4.44	11.56	5.33
LIES	77.56	104.89	84.78	109.44
RIES	59.11	76.56	59.56	81.67
LIES+RIES	136.67	181.44	144.33	191.11
LIPL	26.44	22.22	27.33	20.67
RIPL	19.11	21.89	20.33	20.56
LIPL+RIPL	45.56	44.11	47.67	41.22
LIPS	**54.67**	**76.56**	**61**	**75.22**
RIPS	**37.67**	**64.11**	**48.11**	**72.56**
IPS	**92.33**	**140.67**	**109.11**	**147.78**

Figure 1. Example of visualizations of fMRI data.

Similarly, Figure 2 shows some visualizations used in passive sonar: above, a waterfall diagram shows the angle of various noise sources across the horizontal axis over time across the vertical axis; and below, a table shows the target motion analysis solutions for six different

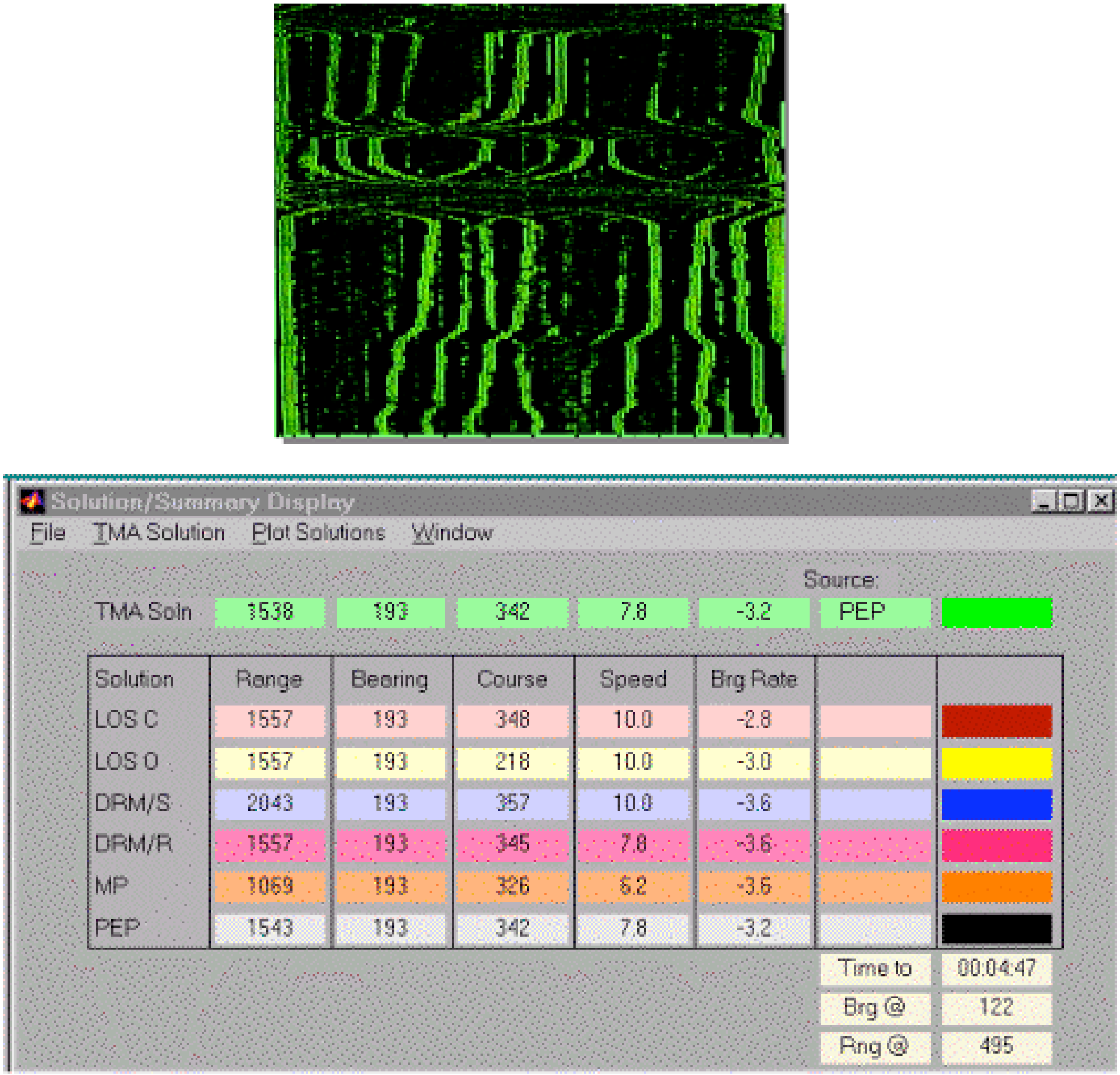


Figure 2: Visualizations used in submarine operations

algorithms. Note the range of the differences in the solutions for the different solutions in the tabular display, and the lack of guidance for interpreting those differences. Of the five displayed variables (Range, Bearing, Course, Speed, and Brg Rate ("Bearing Rate")), only Bearing has the same value in each solution. Moreover, the other values differ greatly. Range, for example, varies from a low value of 1059 to a high of 2043, Course from 210 to 357, and Speed from 62 to 100. Combining these different values into a composite representation of the other submarine compounds those differences, as five completely different scenarios are created. Yet the AO must make a decision about what action to take based on this uncertain information.

Finally, Figure 3 shows a typical meteorological chart, displaying not only land masses, but also multiple weather-related variables in relation to those masses. Variables represented include sea height, wind speed and wind direction. Again, precise values and locations are indicated by wind barbs, alpha-numeric symbols, and defining lines. However, these representations mask a great deal of uncertainty attributable to ranges of values and the dynamic nature of the systems, as well as the questionable trustworthiness and accuracy of the data. Similar mismatches between the implicit certainty of the displayed data and the actual uncertainty of those data exist in the submarine domain, because of both the dynamic nature of the data and the many sources of noise in the ocean environment as the data are collected.

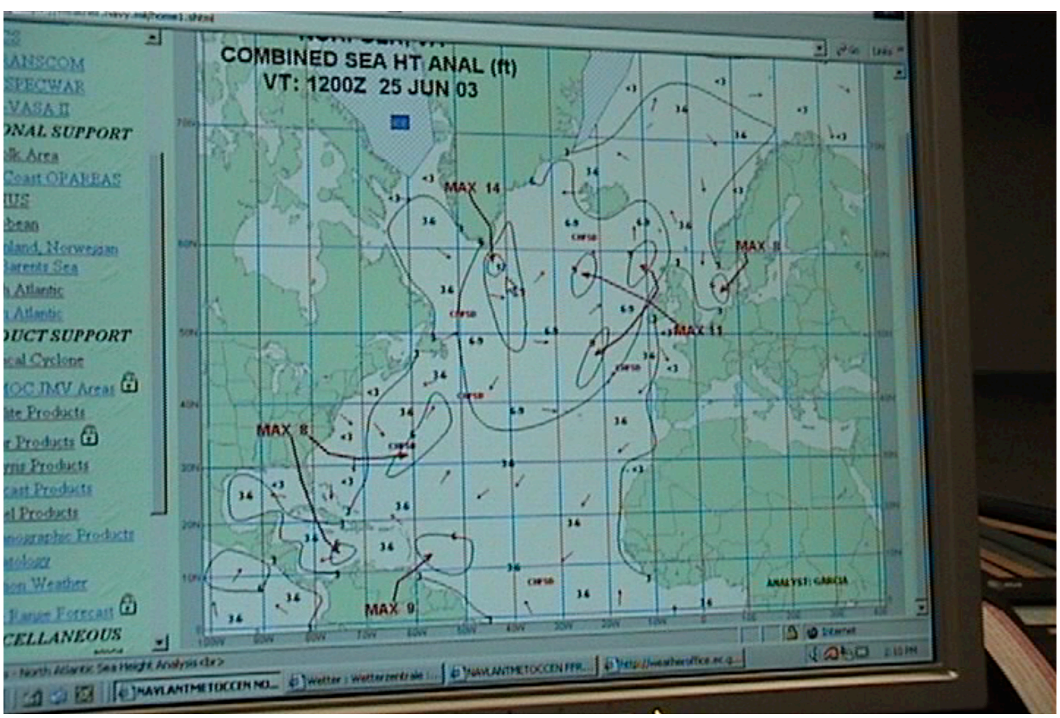


Figure 3. Example of meteorological visualization

The incongruity between the uncertainty inherent in the data and the *lack* of uncertainty explicit in the display presents a serious problem in each of these domains. One of the most time-critical and uncertain events for a submarine is a “close encounter” with another vessel. In such a case the Approach Officer (AO), who is responsible for “fighting the ship” in a hostile encounter, has a very short time in which to assess the evidence (e.g., is the ship friendly or hostile) and take action. As a result, the AO must make safety-critical decisions under extreme uncertainty. Meteorologists must prepare a forecast while contending with inexact information and conflicting model predictions. Weather forecasts are prepared for a customer; inaccurate forecasts can have serious consequences, whether the customer is a Navy pilot who needs detailed flight weather information for the next several hours, a tactical and strategic military planner who needs to know how weather will affect decisions being made for missions, or the general public, who may need to take precautions against severe weather conditions. Results from fMRI and other scientific research are used to inform further research and are often applied to treatments or other problem solutions. fMRI research proceeds along a different timescale from either submarine operations or meteorology; consequently, erroneous results may take weeks, months or even years to be identified and corrected, and may have a significant, negative impact on continuing research.

How do experts think with data and compensate for uncertainty in such uncertain domains? Most studies of decision-making in situations of uncertainty have focused on people’s responses to different gambles based on various probable outcomes. In such studies, the level of risk—or uncertainty—is explicitly manipulated by offering participants the opportunity to win money by placing bets with different probabilities of winning and losing. Extensive work in this area has shown that people respond to this type of uncertainty by relying on heuristics rather than, for example, mathematically calculating the likelihood of gain versus loss (e.g., Tversky & Kahneman, 1974).

However, there are several differences between these tasks and those described above that make it unlikely that similar strategies will be used. First, the stakes are much higher, in that the outcome for the experts has real-world implications, as opposed to simply affecting the results of a laboratory study. Second, the tasks themselves explicitly involve the analysis of large amounts of complex data, some of which is uncertain, rather than a simple choice between two constrained options; furthermore, the goal is not to reach a single decision to take one action rather than another, but rather a deeper understanding of a whole set of circumstances, which may itself inform a decision to be taken later on. Third, the information is presented spatially, rather than mathematically; consequently, it is likely that experts will use a spatial strategy to resolve the uncertainty. Three factors have been identified that play an important role in understanding how experts handle uncertainty: first, complex, spatial domains are rife with uncertainty; second, that uncertainty is often not explicitly displayed; and third, disregarding that uncertainty is likely to have harmful consequences. When uncertainty is not explicitly displayed, the expert must add his or her own understanding of uncertainty to the visualization itself, in order to be able to generate useful solutions to the task at hand.

The challenge for experts working in spatial domains such as those discussed above is thus not only to weigh the implications of explicit uncertainty, such as the potential risks given specific odds, (although this may be a part of the task) In addition, the experts must develop a means of locating areas of uncertainty and re-evaluating the data in accordance with their revised understanding of the likely accuracy of the representation. Consequently, when a visualization displays information that the expert believes to actually be uncertain, in order to use the data, we propose that the expert must modify his or her internal representation to account for that uncertainty. For example, the submarine operator may need to mentally adjust the bearing of a submarine located on radar, in order to represent it as a range rather than an exact angle. In meteorology, the forecaster may need to mentally add information, such as the range of wind speed, or an updated location of a front, to account for likely changes since the data were originally collected and displayed. In fMRI, the researcher may find it necessary to delete an area of activation from his or her mental representation of the data, determining it to be noise rather than viable data. The expert can then use the modified internal representation of the data, which more accurately reflects the external state of affairs, in order to reason and problem-solve about the situation represented in the external visualization.

Constructing and modifying internal representations takes place by mental processes we call "spatial transformations." Spatial transformations are cognitive operations that a person performs on an internal representation (e.g., a mental image) or an external visualization (e.g., a computer-generated image).[1] Sample spatial transformations are creating a mental image, modifying that mental image by adding or deleting features, mental rotation (Shepard & Metzler, 1971), mentally moving an object, animating a static image (Hegarty, 1992), making comparisons between different views (Kosslyn, Sukel, & Bly, 1999; Trafton, Trickett, & Mintz, in press), and anything else a person *mentally* does to a visualization in order to understand it or facilitate problem-solving. Spatial transformations may be used in all types of visual-spatial tasks, and thus represent a general problem-solving strategy in this area.

[1] If the visualization is external, the operation is not literally performed on the external representation, but rather on the internal model of the external representation.

Thus experts viewing unambiguous displays of uncertain data must modify their representation of the data, and spatial transformations are a means by which such mental modifications occur. For example, a meteorologist faced with different weather models making different predictions must somehow resolve those differences in order to construct a single representation of the data. This might be done, for example, by averaging, reconciling, justifying one model over another, creating a composite, or some other means that involves mentally modifying the representation. We hypothesize that when people are more uncertain while working with complex data visualizations, they will perform more spatial transformations than when they are certain. In other words, while spatial transformations may be part of the visual problem-solving toolkit, they are particularly important for resolving uncertainty in complex visual displays.

In order to investigate this hypothesis, we conducted two studies of experts performing their regular tasks. Study 1 presents a re-analysis of previously collected ex-vivo data (Trafton et al., 2000) of a meteorologist preparing a weather brief. It is an initial examination of the relationship between uncertainty and the use of spatial transformations. Study 2 is an *in vivo* study of meteorologists making a forecast and fMRI researchers conducting their own research, designed to elaborate and expand the results of the first study.

Study 1

To explore whether expert meteorologists perform more spatial transformations when uncertain, we re-analyzed one expert forecaster from previously collected data (Trafton et al., 2000). The forecaster was an expert Naval meteorological forecaster with 16 years of forecasting experience; in the past year he had made approximately 600 forecasts. The forecaster worked with a technician and had access to a "regional center," typically staffed with experienced forecasters who are there to provide assistance as well as specialized visualizations.

Procedure

The forecaster's task was to prepare a written brief for an airplane flown from an aircraft carrier to a destination 12 hours in the future (the destination was Whidbey Island, Washington State). The brief was to cover the entire round trip and the forecaster was asked to provide specific weather information for departure, en-route, destination and alternate airfields. In order to do this, the forecaster had to determine detailed qualitative and quantitative information about the weather conditions. This task was a very familiar one for the forecaster. The forecaster was given 2 hours to finish his task, though it took him less than 50 minutes. Further details of the procedure can be found in Trafton et al., 2000.

Coding Scheme

The forecaster's utterances were transcribed and coded using standard protocol analysis techniques (Ericsson & Simon, 1993). We used a purely syntactic approach to coding uncertainty; hedge words like "probably," "sort of," and explicit verbalizations of uncertainty ("We'll have to see if we agree with that or not") were coded as "uncertain utterances." We also extracted approximately 20% of the forecaster's verbalizations that did not have any uncertainty to use as a control. Spatial transformations were coded for each utterance as well. Table 1 shows examples of uncertain/certain utterances as well as spatial transformations.

Utterance	Code	Spatial Transformations (ST)
Nogaps [a mathematical model] has some precipitation over the Vancouver/Canada border (while viewing a visualization)	Certain	No ST
This is valid today	Certain	No ST
Possibly some rain over Port Angeles	Uncertain	No ST

And then uh, at Port Angeles, there's gonna be some rain up at the north, and if that **sort of** sneaks down, we **could see** a little bit of restriction of visibility, but only down to 5 miles at the worst	Uncertain	ST: mentally moving rain [sneaks down]
I **don't think** the uh front's gonna get to Whidbey Island [in 12 hours], but **it should be** sitting **right about** over Port Angeles **right around** 0Z this evening	Uncertain	ST: mentally moving front / animation

Table 1: Examples of certain and uncertain utterances (indications of uncertainty in bold)

Results and Discussion

What is the relationship between uncertainty and spatial cognition? If an expert needs to mentally manipulate a complex visualization in order to understand the uncertainty, we would expect more spatial transformations during the uncertain utterances than the certain utterances. In fact, this is exactly what we found, $\chi^2(1) = 4.1$, $p < .05$. As Figure 4 suggests, when the forecaster was uncertain, he performed about twice as many spatial transformations as when he was certain. Said another way, spatial transformations may be associated in this case with factors other than uncertainty, but about 50% of all the spatial transformations are specifically related to uncertainty.

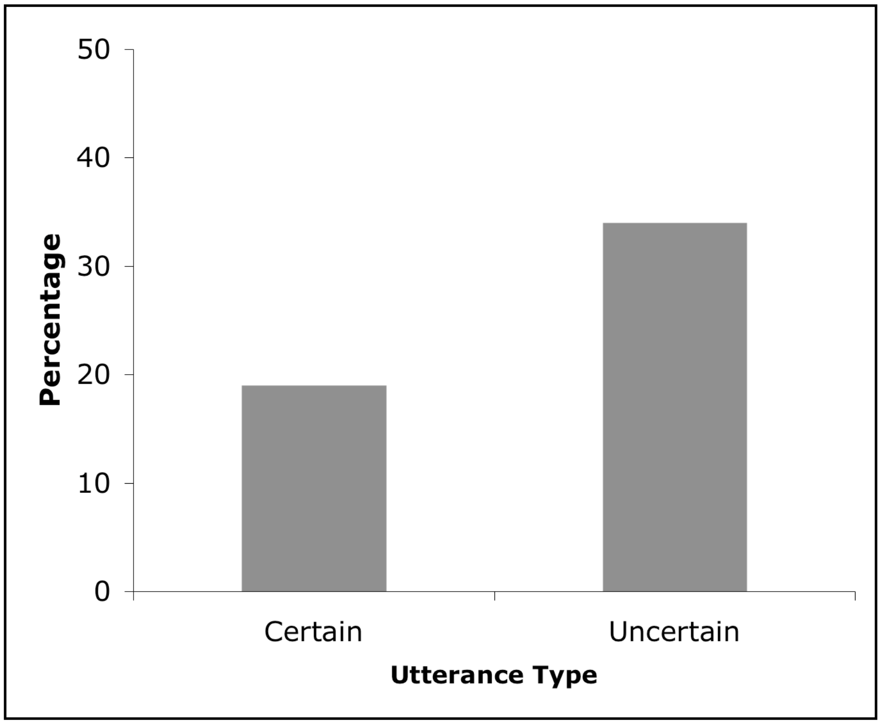


Figure 4: Percentage of utterances containing spatial transformations during certain and uncertain utterances.

Study 2

The results of Study 1 suggest that the forecaster was often mentally manipulating the visualizations in order to understand the uncertainty inherent in the domain. However, there are several obvious shortcomings to this study. First, only one forecaster was examined, and these findings could be idiosyncratic to this forecaster. Second, the grain size (sometimes several complete thoughts) was quite large in the utterances examined and the large size of these utterances may have confounded the coding.[2] Also, the task itself was constructed specifically for the purposes of the experiment. In other words, although the participant was a true expert

[2] This was a deliberate feature of study 1, since it was necessary to see both certainty or lack thereof and the spatial transformation in the same utterance.

performing a task typical of his daily work, the task was not entirely naturalistic, and it is possible that the slightly artificial nature of this experiment affected the experts' behavior and thus their handling of uncertainty. For example, being asked to perform a constructed task might have made the experts somewhat eager to "get the forecast right" and therefore less likely to express their uncertainty.

In addition to these shortcomings, the coding scheme itself may have biased the results. The syntactic nature of the coding scheme for uncertainty captured uncertainty well at a local or immediate level; that is, within the utterance the verbal expressions of uncertainty were most likely an accurate reflection of the forecasters' uncertainty at a given moment about a specific piece of data. However, the coding scheme may not have captured uncertainty at a more global level—for example, if a forecaster was uncertain about the specific location of a front, that uncertainty might have been expressed in only one utterance, but continued in the forecaster's mind for a much longer period. In such a case, the forecaster might gather additional *certain* data in order to try to resolve the uncertainty. Such utterances would be coded as *certain*, when the forecaster's state of mind was, in fact, *uncertain*. Furthermore, the degree of certainty and the use of spatial transformations could not be independently assessed in this dataset.

In order to eliminate these possible sources of bias, and to improve the generalizability of our results, we conducted a second experiment to further investigate the relationship between uncertainty and the use of spatial transformations, and added a second domain. Study 2 was a true *in vivo* study (Dunbar, 1995, 1997), involving both meteorologists and fMRI researchers conducting their own research.

Method

Participants

Participants were two fMRI researchers and two meteorologists. The fMRI researchers had conducted 3 or 4 studies and had an average of approximately 3 years experience in fMRI research. They were thus considered near-experts (see Schunn's chapter in this volume for additional descriptions of these researchers). The meteorologists had many years experience (over 10 years each) working as Navy forecasters, and were thus experts in this domain.

Procedure

The experiment took place at the participant's regular work location, and all participants had access to all the tools, visualizations and computer equipment that they usually employed. All participants agreed to be videotaped during the session. Participants working alone were trained to give talk-aloud verbal protocols (Ericsson & Simon, 1993). All participants were instructed to carry out their work as though no camera were present and without explanation to the experimenter. It is important to emphasize that all participants were performing their usual tasks in the manner in which they typically did so, without interruption from the experimenter.

While the participants performed the task, the experimenter made note of "interesting events". Interesting events consisted of any event that seemed to pique the participant's interest, major changes in the computer display, such as a new visualization or application, an event that spurred a burst of participant activity, and the like. In other words, "interesting events" were those that struck the experimenter, in this on-line coding, as non-routine and worthy of further probing.

After the task was completed, the experimenter showed the participant a one-minute segment of the video surrounding the "interesting event"; we shall refer to these video segments as *interesting minutes*. For each interesting minute, after reviewing the videotape, the experimenter asked the participant “What did you know and what did you not know at this point?” Participants’ responses to these questions were also recorded on videotape.

Coding

All utterances, from both the *in vivo* data and the interview data, were transcribed and segmented according to complete thought. For the *in vivo* data, all spatial transformations for the interesting minutes were identified, as described in Study 1. For the interview data, a second, independent coder (from a different lab) coded each utterance as *certain* or *uncertain*, using the same criteria as in Study 1. The difference between the uncertainty coding for the two experiments was that in Study 1, the coding scheme was applied to the *in vivo* data, whereas in Study 2, it was applied to the interview data. Based on the percentage of uncertain utterances in each interview minute, the corresponding *in vivo* minute was coded as *certain* (when fewer than 10% of the utterances were uncertain), *mixed* (when 10 to 20% of the utterances were uncertain), or *uncertain* (when more than 20% of the utterances were uncertain).

The participants’ retrospective utterances about their task performance provide an independent measure of their uncertainty during problem-solving, and thus address the concern discussed above about possible bias in the uncertainty coding in the first study. It is important to emphasize that the coder for the *in vivo* data did not have access to the interview minutes, or to any coding associated with those minutes. Likewise, the coder of the interview minutes did not have access to the *in vivo* data or the spatial transformation coding associated with it. Thus the two coding schemes were applied completely independently of one another, further precluding possible coding bias in the association between uncertainty and spatial transformations.

Results and Discussion

Quantitative Analysis

Of the 19 interesting minutes, 5 were coded as certain, 3 as mixed, and 11 as uncertain, thus confirming the large amount of uncertainty practitioners in these domains must contend with. Participants used least spatial transformations in the certain minutes, more in the mixed minutes, and most in the uncertain minutes, and this difference was significant, $\chi^2(1) = 20.85$, $p < .01$. Figure 5 clearly shows the increase in the use of spatial transformations that accompanies the shift from greater certainty to greater uncertainty. These results support our basic hypothesis, that people will use more spatial transformations when they are uncertain than when they are certain.

One possible explanation for this increased use of spatial transformations is that it is part of a pattern of generally increased activity that occurred during uncertainty. That is, perhaps the participants were simply doing more things, or thinking more, when they were uncertain than when they were certain. In order to test this possibility, we examined the number of interface actions the scientists took in the certain and uncertain minutes. Interface actions were defined as manipulations of displayed data, and included closing visualizations and opening new ones, adjusting images (e.g., zooming in to enlarge them), opening and closing windows, etc. in order to advance an understanding of the data. As Figure 6 shows, the number of interface actions was *not* related to the certainty/uncertainty coding. Participants used about the same number of interface actions (1.2 compared with 1.8) when they were uncertain as when they were certain, with a slightly increased use of spatial transformations when their uncertainty level was mixed. The lack of obvious linear increase shown in Figure 6 strongly suggests that these practitioners' general level of activity was independent of their level of uncertainty. It certainly indicates that there is no reason to suspect that participants were conducting more interface actions when uncertain than when they were certain.

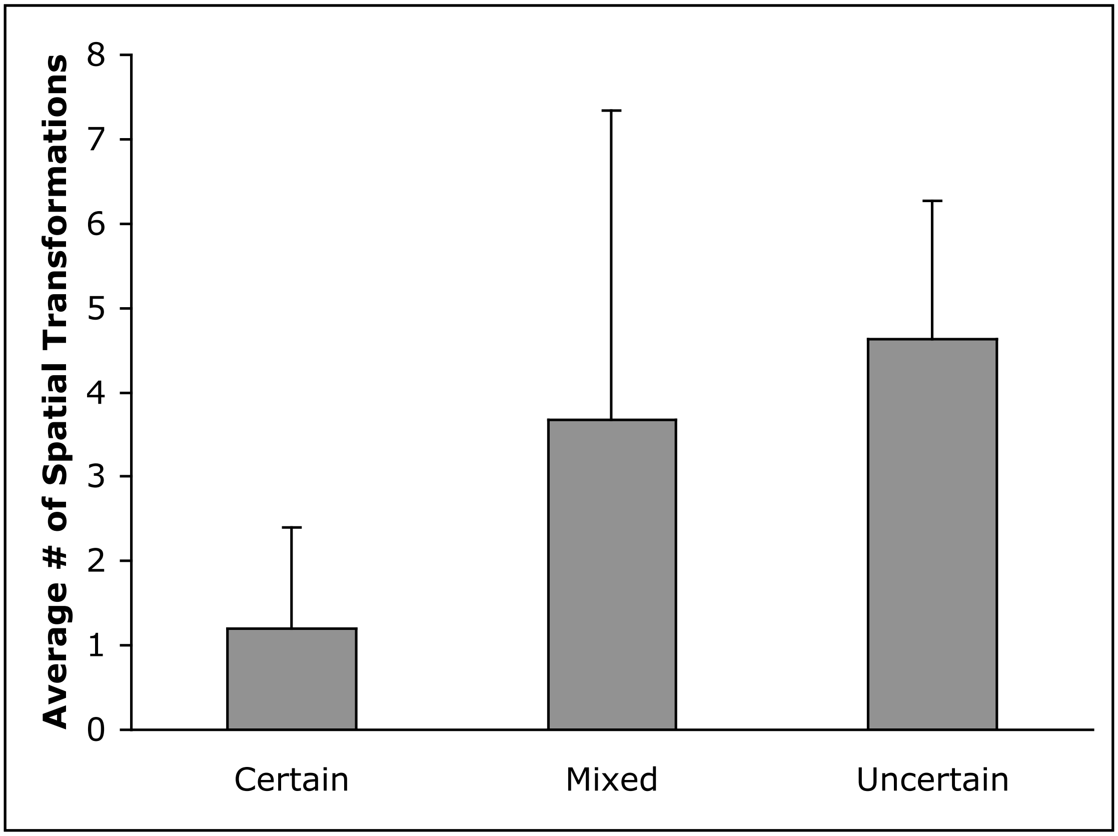


Figure 5. Average numbers of spatial transformations used in certain, mixed, and uncertain interesting minutes

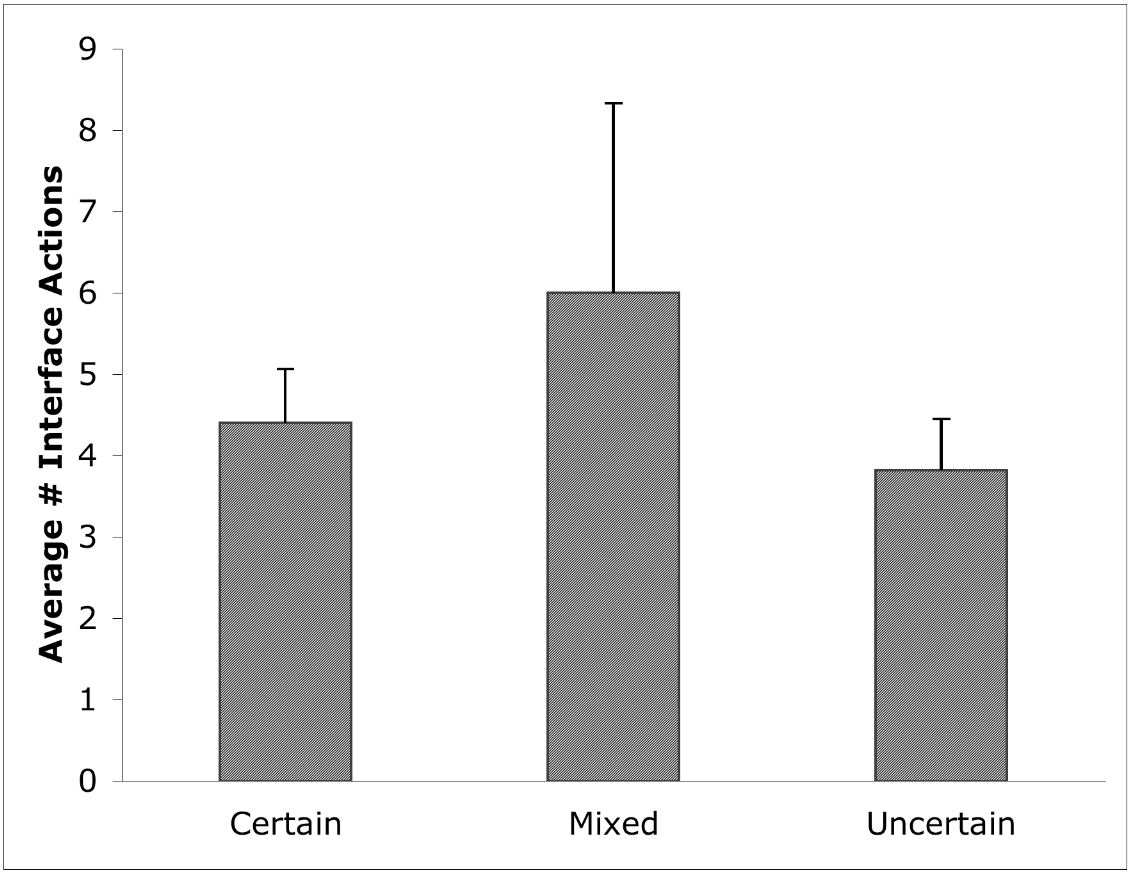


Figure 6. Average number of interface actions in certain, mixed, and uncertain interesting minutes

Qualitative Analysis

We hypothesized that people would use more spatial transformations when they were uncertain than when they were certain, and this hypothesis was supported. We further proposed that spatial transformations would provide a means for practitioners to project their own knowledge—or suspicions—of uncertainty onto the external display, thereby constructing a more accurate internal representation of the data. Experts would then be able to work with this projected, internal representation of uncertainty in order to carry out the task at hand.

There are many instances of this type of process in these data. For example, consider the following extract from the protocol of one meteorologist using the visualization in Figure 3:

> *You also have a 12 max 14—winds are not supporting that. The next chart has it moving down further to the south. There is a low coming off the coast that is probably getting around, so I would move it further to the south. And that just supports what I said about ours, OK.*

The meteorologist's overall goal in this section was to determine whether or not a high seas warning is warranted. First, the meteorologist extracts information about the wave height from the chart (*You also have a 12 max 14)*, but then she realizes that there is a mismatch between this information and another source of relevant information, wind speed and direction (*winds are not supporting that)*. (The meteorologist had earlier commented that what drives the seas is the winds.) This conflict between two sources of information gives rise to uncertainty on her part. She consults a third source of information in memory—a chart she had looked at previously—and comparing this information with the location of the specific wave height in the current visualization realizes the charts are suggesting two different locations for the "12 max 14." Which location should she use to inform her forecast? With all this information in mind, the forecaster then introduces a fourth data source, an area of low pressure. The interesting thing about this low is that it is not represented on the visualization. In other words, the meteorologist first mentally creates the low (*there is a low coming off the coast)* and then—also mentally—moves it around to position it in relation to the other data already in mind (*that is probably getting around)*. She then uses the implications of this new internal representation to mentally relocate the area of "12 max 14" wave height *(so I would move it further to the south)*. Inspecting this revision, she is able to resolve the uncertainty *(And that just supports what I said about ours, OK)*, because, the more southerly location makes better sense given the spatial transformations she has performed. Spatial transformations thus allow the meteorologist to confirm the uncertainty (comparison between currently displayed and remembered data), to introduce relevant information that is not displayed (the low), to manipulate that information in relation to the displayed data (moving the low), to use domain knowledge to project the implications of these manipulations (relocate the "12 max 14"), and finally to evaluate the results.

Similar examples of the way spatial transformations function are found in the fMRI data. At one point, one of the fMRI researchers expressed some uncertainty about the pattern of activation on the display:

> *They're all decreasing below baseline, but it's still hard to know what decreases in activation mean, so I don't know what's going on there. I'll see if Jane knows. I don't think anybody really knows what decreases really mean.…So punish looks like it's a little bit higher, but that's probably not significant, and it looks like* [these two] *are the same.* [Pause] *That is—ah! If that's really postcentral gyrus that would make sense—they're hitting a button in both cases.*

This entire episode takes place in a context of uncertainty. The researcher begins by reading off information (*They're all decreasing below baseline*), but then acknowledges that that information isn't especially useful in terms of understanding brain functioning, because it is uninterpretable (*I don't think anybody really knows what decreases really mean).* He continues to try to read off information, comparing activation in two conditions (*punish looks like it's a little bit higher)* and interpreting what it might mean *(but that's probably not significant).* Unfortunately, the experiment doesn't seem to have worked (*it looks like these two are the same*). He seems to feel a significant amount of confusion about what the data are saying, because this was an experimental manipulation in which significant differences in activation between conditions were predicted. Then he appears to have an insight (*that is—ah!)* and proceeds to perform a spatial transformation of the data, by positing the area of activation is actually in an adjacent region of the brain *(if that's really post-central gyrus, that would make sense*). In other words, if the display is not really showing what it purports to, but the area of activation is, in fact, the post-central gyrus, the pattern of activity "would make sense." The researcher further projects the participants' actions onto the representation (*they're hitting a button in both cases)*, a supposition which would account for the lack of difference between the conditions. The uncertainty and confusion can thus be resolved

These two examples are typical of the kind of problem-solving behavior demonstrated during the uncertain minutes: the experts used the visualizations as an initial source of data but then mentally manipulated the visualization in order to accommodate and resolve their uncertainty about what the data really represented. Comparing this behavior with their use of the visualizations during the certain minutes further highlights the role of spatial transformations during uncertainty. Consider this excerpt from one of the certain minutes in the fMRI domain:

> *Now we're going to do a contrast of areas that are active for words but have, really compare the contrast between the two. Ah, there we go. Now that's what I like to see; it makes a lot more sense. You can see, this is, this is, um, this is beautiful. This is exactly what you want to see for this type of data. You see a trend going right up the um, right here, [these] coordinates—left is right*[3]*—so right along the right ventral visual pathway, you see this nice stream of activation.*

In contrast to the two examples discussed earlier, in this instance the researcher's expectations about differences in activation patterns for the experimental conditions are clearly met. He announces the particular comparison he is going to make, and the data are displayed as he has predicted. His utterances consist entirely of either reading off information from the display (e.g.,

[3] The comment "left is right" is not a spatial transformation but rather refers to the fact that the left side of the graph represents the right side of the brain; in other words, it describes the mapping of graph to data.

You see a trend going right up the um, right here; right along the right ventral visual pathway, you see this nice stream of activation) or of exclamations of satisfaction (e.g., *this is beautiful; it makes a lot more sense*). There is no uncertainty in his interpretation of the data, and there are no spatial transformations. Perhaps because his expectations have been met, he has no reason to be anything other than certain about the accuracy of the display.

General Discussion and Conclusion

We conducted two studies to investigate our hypothesis that people using complex visualizations would use more spatial transformations when they are uncertain about the visualization than when they are certain. The results of both studies support this hypothesis. Study 1 provided an in-depth examination of the protocol of a single expert preparing a weather brief, using an on-line measure of uncertainty. Study 2 provided an extension of study 1 by expanding the number of domains and participants, and by developing independent coding schemes for uncertainty and spatial transformations. Furthermore, the results of study 2 showed that participants were not merely engaged in greater general problem-solving activity as uncertainty increased, as the number of interface actions was unrelated to their level of uncertainty.

One caveat to the results of our two studies is that the data are correlational, and consequently, we are unable at this time to conclude that uncertainty causes the experts to use spatial transformations. Nonetheless, given their contrasting behavior when participants were more certain, we do believe that the use of spatial transformations is a strategy by which experts deal with, and in some cases resolve, uncertainty in spatial domains. Although it is possible that a spatial strategy could be used in non-spatial domains, such as the likely responses to different gambles based on various probable outcomes discussed earlier in this paper, we think it unlikely, in fact, that this is the case. Our main point is that in spatial domains, when experts must handle uncertainty, they use spatial transformations to do so.

Specifically, we propose that because the participants were experts or near-experts, they had significant domain knowledge as well as the ability to recognize a large number of patterns in their given domain. Thus, they are aware (through training) that whereas the data appear certain, there are in fact many reasons to doubt its seeming precision and lack of ambiguity. Furthermore, their pattern-recognition capacities (developed from experience) can help guide them to identify those instances when they should doubt the data and make necessary modifications. They make these modifications by means of spatial transformations (mental manipulations) of the displayed data, by mentally creating or deleting data points, mentally moving objects around, mentally animating data, mentally projecting stored memories of past experiences, and the like. The modified representation—now internal—becomes a new resource with which to reason about the data. This internal representation can itself be further modified by additional spatial transformations, as new information is obtained. Thus, in long chains of reasoning, a complex dialogue between the external and internal visualizations may evolve, in which each can be updated until the expert is satisfied that he or she has enough information to reach a conclusion or to make a decision. In contrast, when they are certain about the data, experts use only the external visualization as a source of information, focusing their problem-solving activity on reading off data values or other relevant types of information. When the data are considered reliable and sufficient, no further action is required.

How do our results mesh with the results of other research presented in this volume? Several researchers address issues of reasoning about uncertainty; however, the majority of those papers concentrate on uncertainty as the statistical notion of variability and error. In contrast, the research presented in our paper focuses primarily on reasoning with visual data, rather than quantitative or statistical data. Whereas it is true that participants in these other studies used or even created graphical representations of statistical uncertainty, they did so mostly in order to represent, manipulate, and understand numerical variability. This use is consistent with the general use of graphs in modern statistics as a tool to understand numbers rather than as a means to examine real underlying effects in the data.

Although variability and error are indeed factors for our experts, there are many more varied sources of uncertainty in the data that they must address—visual uncertainty (e.g., "I'm not sure if my area is being masked by this whole temporal area"), uncertainty about whether the data are complete or accurate (e.g., "I would move [the low] further to the south"), uncertainty about whether the data are outdated (e.g., "I'm somehow having to run off an old model, which is frustrating"), interpretive uncertainty ("It's still hard to know what decreases in activation mean…I don't think anyone really knows what decreases in activation mean"), and even uncertainty about whether the data in question represent the correct view (e.g., "…[he made] design files, but I think I told him to do it the wrong way")—to name a few.

The different areas of focus (visual versus quantitative) raise important questions about the reasons for that divergence. Does the difference arise simply from what the individual researchers chose to examine, or does it reflect a genuine difference in an understanding of what data analysis is really about? In other words, to what extent does data analysis, in the sense of interpreting statistical concepts, capture the nature of "thinking with data"?

Statistical reasoning is, in fact, a crucial component of thinking with data. For many people, it is also a difficult process because it requires an understanding of concepts of uncertainty (variability and error), in order for useful interpretations to be constructed. However, we propose that statistical representations of uncertainty are just one of many aspects of uncertainty in thinking with data, and that the full complexity of what constitutes uncertainty when thinking with data might be masked in some experimental settings.

One important factor that may contribute some insight to this issue is the "in vivo" nature of our study. Not only were our participants asked to do tasks that they regularly perform, but they also had access to a wide range of data and analysis tools. Their problem-solving goal was internally motivated. In contrast, the other studies were conducted in a classroom setting with its attendant requirements to complete certain structured learning objectives prior to moving onto another unit of study. The participants were also engaged in problem-solving activities, but their motivation may have been external rather than internal. "Thinking with data" may necessarily mean different things in these different settings, one naturalistic and the other instructional.

A second difference in our study that may be relevant is the fact that our participants were experts with many years of experience and a great deal of accumulated domain knowledge, in contrast to either children learning both the content and the methodology or college students working on an abstract task. As we noted above, we believe that an integral part of our participants' response to uncertainty was initially *recognizing* that uncertainty existed in the data. The large number of patterns stored in memory from prior experience, and the ability to recognize and interpret them, were instrumental in discerning uncertainty, even when it was not explicitly portrayed. Novices and those working in abstract domains cannot rely on pattern recognition mechanisms to help identify areas of uncertainty; nor do they necessarily have the requisite training to understand at a general level the uncertainty inherent in any empirical data (due to measurement error, problems with experimental design, and so on). Thus, it is not surprising that in addition to performing more complex tasks, the experts were more aware of the potential for multiple sources of uncertainty in the data, many of which were not explicitly represented. We suggest that this awareness and ability to exploit it are important factors, specifically in dealing with uncertainty, and more generally in thinking with data. However, how such awareness develops remains an open question.

The issues raised by the differences in these two approaches to data analysis are important as we think more broadly about what it means to "think with data." Are the processes the same regardless of the situation? Or does this, in fact, represent two qualitatively different tasks for experts and novices, or in real-world science as opposed to a laboratory or instructional setting, or in formal science rather than everyday or informal reasoning? A next step to investigate these issues further would be to study how novices in a domain (with some years of formal training and thus the requisite domain knowledge) handle uncertainty.

References

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 365-395). Cambridge, MA: MIT Press.

Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T. B. Ward & S. M. Smith (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 461-493). Washington, DC, USA: American Psychological Association.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. (2nd ed.). Cambridge, MA: MIT Press.

Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*(5), 1084-1102.

Kosslyn, S. M., Sukel, K. E., & Bly, B. M. (1999). Squinting with the mind's eye: Effects of stimulus resolution on imaginal and perceptual comparisons. *Memory and Cognition, 27*(2), 276-287.

Shepard, R., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*, 701-703.

Trafton, J. G., Kirschenbaum, S. S., Tsui, T. L., Miyamoto, R. T., Ballas, J. A., & Raymond, P. D. (2000). Turning pictures into numbers: Extracting and generating information from complex visualizations. *International Journal of Human Computer Studies, 53*(5), 827-850.

Trafton, J. G., Trickett, S. B., & Mintz, F. E. (in press). Overlaying images: Spatial transformations of complex visualizations. *Foundations of Science*.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.