Detecting and Resolving Informational Uncertainty in Complex Domains

Christian D. Schunn (schunn@pitt.edu), Tsunhin Wong (tjw23@pitt.edu),
Waleed Manzoul (wam8@pitt.edu), Jamie Kamer (jrk17@pitt.edu)
LRDC, University of Pittsburgh, 3939 O'Hara St.
Pittsburgh, PA 15260 USA

Julie Harris (JHarris@Chatham.edu)
Department of Psychology, Chatham College
Pittsburgh, PA 15260 USA

J. Greg Trafton (trafton@itd.nrl.navy.mil), Susan B. Trickett (stricket@gmu.edu)
Naval Research Lab, Code 5515
Washington, DC 20375-5337 USA

Abstract

A common feature of problem solving in real world complex domains is detecting moments of high informational uncertainty and trying to resolve this informational uncertainty. Yet, relatively little is known about these important aspects of real world cognition. Expert and novice performance is contrasted for relative use of informational uncertainty indicator strategies and problem solving strategies across two very different complex domains: weather forecasting and geoscience planning and data analysis. Strong differences across domains were found for problem solving but not indicator strategies. The differences between domains were exaggerated in experts. Indicators strategies generally appeared associated with spatial representations, whereas only a few of the uncertainty problem solving strategies did so.

Keywords: Ill-defined problem solving; uncertainty; expertise; spatial reasoning

Introduction

A well-worn path in cognitive science is the distinction between well-defined and ill-defined problems (Newell & Simon, 1972). And yet, a clear definition of "ill-defined" remains an illusive anti-prototype definition: problems that are NOT well-defined. Problems can be ill-defined in a three classic ways: lacking a clear start, lacking a clear end, and lacking a clear sense of allowable actions along the way from start to goal. Yet, definitions by absence of features are not very specific. Cognitive science of problem solving could explore the endless variations of ways of NOT being well-defined on each of those three dimensions. But to make more rapid progress, one heuristic can be to turn to the world around us, and explore what kinds of ill-defined tasks regularly occur, to focus research attention on that range of ill-defined tasks first.

We have been looking at experts in a variety of highly ambiguous yet regularly solvable tasks: submariners trying to localize another submarine with passive sonar, weather forecasters making a weather forecast, scientists analyzing data and planning experiments, and engineers designing innovative new products (e.g., Trafton et al., 2006; Christensen & Schunn, in press). For experts, each of these domains is primarily well-defined in the classic sense: the start state is well understood, the set of typical operators are well honed, and the goal state is usually clear (although there may be more than one possible acceptable goal state). And yet, even for experts, these domains included a very high level of ambiguity. Problem solving in these domains is more heuristic than algorithmic, even for experts.

A common feature across problem solving in all of these domains that is an important source (and sometimes THE important source) of ambiguity is informational uncertainty: psychological uncertainty about the accuracy of information about the current state of the world. That is, unlike in classic well-defined problems, the information that defines the current state is interpretable in many different ways, even by experts. The sources of the informational ambiguity can include physical factors (such as measurement error of a device), computational factors (such as statistical artifacts of smoothing or averaging algorithms), visual input factors (such as graphical information displays including only some of the relevant dimensions), and cognitive factors (such as prior procedural errors or memory retrieval errors). In these domains, a core element of expertise is being able to detect states with high levels of informational uncertainty and being able to resolve that informational uncertainty through problem solving. Informational uncertainty is contrasted with decision-making uncertainty, which adds uncertainty about the consequences of following different choices.

In this paper, we explore in depth the particular indicators of uncertainty that problem solvers tend use to detect informational uncertainty and the particular strategies that they tend use to resolve uncertainty. Is resolving uncertainty primarily a mental reasoning task, using inductive and deductive reasoning, as one believe from the focus of most texts on higher-level cognition? If it is mental reasoning, is the reasoning primarily verbal, or is spatial reasoning a core component (Trafton et al., 2006)? Or is the problem solving more heavily dependent upon external representations (Hutchins, 1995), relying on epistemic markup of visualizations (Kirsch & Maglio, 1994), and/or iterative search through externally stored data?

We use a mixture of verbal protocols and spatial gesture analysis to examine whether the elements are primarily involving verbal internal representations or spatial internal representations. Although many complex domains have rich visual-spatial external input, problem solvers can re-represent information into a variety of forms. To understand the nature of cognitive processing at a detailed level, verbal/spatial is generally thought of as an important first cut at unpacking processing.

We further use an expert/novice contrast to explore which elements require more experience. One naïve assumption might be that novices are just generally uncertain, and do not have particular strategies for noting cases of informational uncertainty. More likely is the possibility that experts make more use of knowledge-rich methods of detecting and resolving informational uncertainty. But it is also possible that experts have discovered some domain-specific strategies (but not requiring much knowledge) for resolving uncertainty that novices do not have at all, or do not tend to use very often (Siegler, 1996).

Understanding problem solving in just one domain requires a lot of effort. The result can be quite rich descriptively, but the generality can also be suspect. Our approach is to analyze data from two different domains. One should not really think of this contrast between domains as a single unconfounded variable contrast because any two domains have many features that are different. Instead, the two domains should be thought of as two cases for exploring the generality of the obtained picture of problem solving with informational uncertainty.

Methods

Domains

MER The first domain is remote planetary science. Specifically, it involves data collected at JPL in Pasedena, CA of scientists analyzing and planning during the first 90 days of the Mars Exploration Rover (MER) mission, in which two robots crawled around Mars on the basis of commands sent daily, and sent back images and other instrument data via satellite relay also on a daily basis. The scientific goals of the mission roughly involved characterizing the geology, geochemistry, and atmospheric conditions now and in the past on Mars. In this domain, the time pressure for resolving informational uncertainty is low (can take months in some cases), and the overall strength of prior expectations is low given that this level of detail of data from another planet was not previously collected.

METOC The second domain is Meteorological and Oceanographic forecasting (METOC). Specifically, it involved weather forecasters making weather forecasts for a particular point in time (typically 1 to 7 days in advance) at a particular location (typically not the forecasters exact location). In this domain, the time pressure for resolving informational uncertainty is moderate (should be resolved in an hour or so), and the overall strength of prior expectations

is high given that the participants have made such forecasts in very similar locations many times before.

Participants

MER The participants were senior and junior scientists in the domains of geology, geochemistry, soil science, and atmospheric science. Some were employees of the Jet Propulsion Lab, but others were visiting from various universities and research centers for the duration of the nominal 90-day mission. On any given day, approximately 20 scientists (divided into 5 thematic groups) worked on analyzing data and planning the next day's activities for a given Rover. During the day, scientists attended three planned full-group meetings, but primarily worked informally in various subgroup sizes.

Expertise was coded into approximate expert/novice categories on the basis of age appearance. Unlike some other disciplines (e.g., education), researchers in this area are rarely second career, and thus age is a reasonable proxy of expertise here. Given the very large range (20s to 60s), this coding was relatively easy to do. Those scientists in the their 40s and older were considered experts, and those scientists in the 30s and younger were considered novices.

METOC The 17 novice participants were 3rd and 4th year weather forecasting students at a college that offered a degree in weather forecasting. The 6 expert participants were military weather forecasters each with over 10 years of weather forecasting experience. Note that in both domains, the expert groups are beyond the 10-year point associated with high levels of expertise. Also note that in neither domain are the novices uninformed initiates. Instead, the novices understand the basics of the domain, have strong identity with the domain, but simply have noticeably less experience in the domain than the experts.

Data Collection

MER Video cameras were setup in various locations around the room on a rotational basis, on tripods at a height of approximate 2m, located slightly above 60in touch-screen displays. The cameras were frequently in the space, and were turned on before scientists entered the space in the room according to a sampling schedule that covered early, middle, and late days on the mission, producing approximately 400 hours of video. Although the scientists had signed informed consent forms and were generically aware of the video recording, the location of the cameras and their omnipresence in the space meant that participants were generally unaware of the camera. Fourty snippets of 5 to 30 minutes of video were sampled to be coded on the basis of: 1) two or more individuals having an informal conversation about data analysis or mission planning; 2) the conversation takes place on camera; and 3) no nearby individuals were having another conversation that interfered with the audio capture. The conversations were transcribed verbatim and segmented at the level of complete thought, producing 5878 segments of speech.

METOC The participants were asked to produce weather forecasts as they normally would, but while providing a think aloud protocol (Ericsson & Simon, 1993). The participants worked on a computer, accessing websites and other weather products as they normally would, or using equivalent printouts displayed on a wall—both modes are common for weather forecasters. The camera was located behind them, capturing work on the computer or with the wall maps over the shoulder. The think-alouds were transcribed verbatim and segmented at the level of complete thought, producing 11,582 segments of speech.

Coding Details

From the verbal data, significant moments of uncertainty were found. These significant moments were defined as moments during which the problem solver explicitly expressed some uncertainty about current information and proceeded to do some external, mental, or social problem solving to resolve it. This coding also noted when the moment began, ended, and when the problem solver digressed off-topic (those segments were excluded). A primary indicator was coded for each significant moment (i.e., what led to the detection of that bit of uncertainty?). Each problem-solving step was separately coded for problem solving strategy, because problem solving in a significant moment typically involved many different strategies to obtain resolution. There were 31 such moments in MER and 82 in METOC.

Indicator Taxonomy Four indicators of informational uncertainty were created in a bottom-up fashion from the data. They are listed in Table 1.

Table 1: Taxonomy of indicators used by participants to detect informational uncertainty.

Indicator	Definition
No Pattern	Speaker claims that they do not see a pattern — no possible alternatives are given.
Ambiguous Pattern	Speaker is unable to distinguish which pattern is present; particular alternatives are listed
Atypical Pattern	Speaker claims that pattern present is not the normal pattern for a particular situation
Inconsistent Pattern	Speaker claims that data from two different sources are contradictory

Uncertainty Problem Solving Taxonomy Thirteen codes were developed in an iterative, bottom-up fashion from the data to reflect the problem solving steps used by participants to address informational uncertainty, as well as our ability to reliably code distinctions. Participants would frequently use several strategies in resolving a given moment of uncertainty. The strategies listed below planning are ones more tightly associated with the final resolution of the uncertainty, whereas the ones above tend to occur earlier during the problem solving.

Table 2: Taxonomy of problem solving strategies used by participants to resolve informational uncertainty.

STRATEGY	DEFINITION
Mark uncertainty	Explicitly declaring a case unresolvable for now
Check likely errors	Recheck steps for technology or human error
Domain knowledge usage	The use of domain-related knowledge to reason
Mental spatial transformation	Reorienting / moving / molding information mentally
Physical Spatial Transformation	Physical manipulation of real or computer objects
Detailed view / more data	Detailed examination of current or new data
Planning	All future-oriented outlines of (non-immediate) actions, whether they are definite, entertained, or conditional
Focus on reliability of sources	Use a source among multiple inconsistent sources—must give explicit reasoning for picking one source over another.
Discount data	Declare data as unreliable or problematic—must give a reason
Adjust for known deviations	Adjust current image for general deviation/irregularity
Average across sources	Take the middle value between two or more data sources
Bound uncertainty	Declare that an instance of uncertainty is within a finite range in scale or within a finite set of possibilities
Unsupported resolution	Resolution made but no reason given

Spatial Gesture Coding In the METOC case, it was impossible to observe participants' gestures given the location of the camera and the restricting effect of interacting heavily with the computer. However, gestures in the MER domain occurred frequently and were generally visible to the camera. These gestures were coded for spatial content, using a coding scheme developed and validated elsewhere (Schunn et al, 2007). In brief, the scheme removes non-spatial or ambiguously-spatial gestures such as beating, deictic (pointing to an object), iconic (hands represent a categorical physical object), and metaphorical (hands use space to represent a non-spatial quantity like time) gestures. The spatial gestures are coded for the broad spatial nature, specifically whether they are 2D display-based gestures, small 3D gestures, or large 3D gestures. In this context, the great majority of gestures were small 3D gestures, and thus we simply examine whether a spatial gesture occurred or not.

Reliability All coders were trained to criterion for each coding dimension. Two separate coders exhaustively coded

all data on all dimensions, and all discrepancies between coders were resolved through discussion. All dimensions had at least moderate Cohen's Kappas (i.e., K>.4) for interrater reliability. It was to increase the <u>effective</u> reliability that each segment was double-coded.

Analyses We use X^2 tests of independence between various dimensions to measure the degree of associate between variables (e.g., group and strategy), collapsing data across individuals, and dropping rows with expected cell values below 5. The significant moments do not occur often enough per individual to establish reliable measures at the per individual level. However, we have examined the data at the individual level to make sure the phenomena are not driven by the behavior of just a few participants. All figures include SE bars to provide an approximate sense of which pair-wise differences are statistically significant.

Results

Indicators of Informational Uncertainty

What indicators were commonly used in each of the two domains? Figure 1 shows that: 1) all strategies were used in both domains, 2) No Pattern is rarely used to identify significant moments of uncertainty, and 3) that the two contexts had considerable variability in strategy use, $X^2(2)=8.84$, $p<.02$. In the highly novel situation with a slow data acquisition rate, ambiguity of information was the most common source. In the more common situation with high data acquisition from multiple sources, inconsistent information was the most common source.
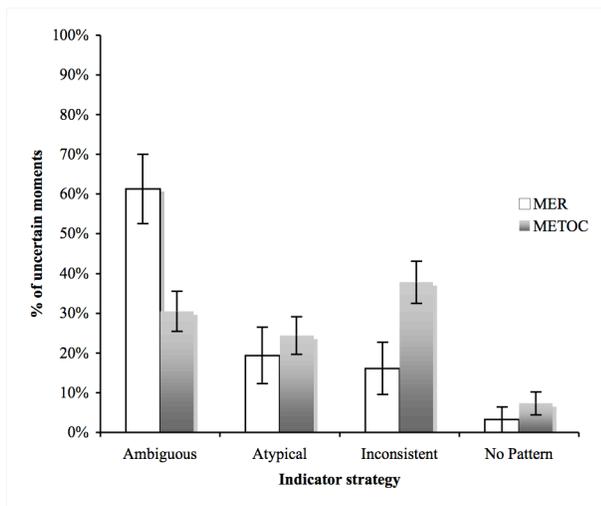

Figure 1: Indicator strategy use by domain.

Figure 2 show expert/novice differences in the reliance on particular uncertainty indicators within each domain. In both domains, the distribution of indicator use between experts and novices was quite similar (within the statistical resolution of the data). In other words, expertise did not influence how uncertainty was detected, either at the level of which strategies were used at all, or roughly how often

each strategy was used, at least over the range of expertise examined here.
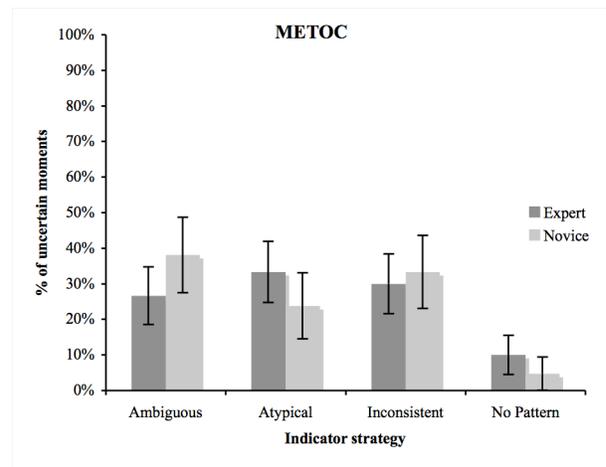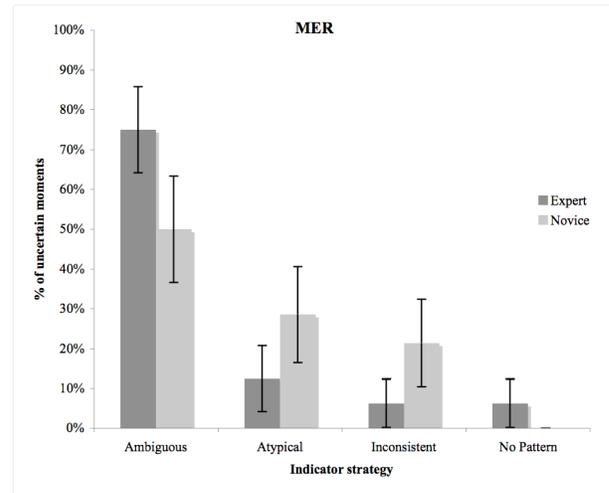

Figure 2: Expert/Novice relative use of informational uncertainty indicator strategies within each domain.

To understand the role of spatial mental representations in uncertainty detection, we examined the percentage of indicator segments overlapped with spatial gestures in the MER domain. Overall, indicator segments contained spatial gestures (of any kind, although almost exclusively small, 3d gestures) 26% of the time, well above the overall rate 11% at which speech segments had spatial gestures (p<.01). Moreover, the three regularly occurring indicators all had approximately equally high rates of co-occurring spatial gestures (varying between 20% to 33%).

Problem Solving Strategies

Overall, all strategies did occur in the data. The most common strategies were: More data, Domain knowledge usage, Adjust for known deviations, Planning, Mental spatial transformation, and Focus on reliability of sources. Several strategies were used rarely (in less than 3% of problem solving steps): Bound uncertainty, Average across sources, Check likely errors, Discount data, Mark uncertainty, Physical Spatial Transformation.

The relative use of all of the more common strategies did differ by domain, $X^2(5)=218.23$, $p<.0001$. Figure 3 shows the frequency of problem solving strategy use in each domain for the six most common strategies. In METOC, there was relatively more use of examining additional data, adjusting for known deviations, using mental spatial transformations, and focusing on the reliability of sources. By contrast, in MER, there was relatively more use of domain knowledge and planning. While the differences might reflect the individual vs. group nature of the data collection methods, the differences are easily understandable in terms of the nature of the two domains. In METOC, many different data sources for a given problem exist (e.g., different weather models, close time points, related dimensions), making for easy examination of additional data and allowing for the heuristic of using reliability of different sources. In MER, data trickles in at a slower rate, and thus working from first-principles domain knowledge and planning additional data collection are more relevant operators for resolving informational uncertainty.
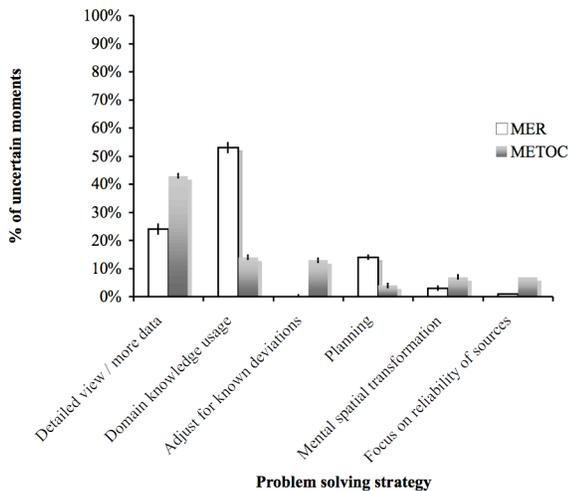

Figure 3: Relative frequency of the most common uncertainty problem solving strategies by domain.

Within each domain, there were also small but statistically significant effects of expertise levels (see Figure 4). In the MER domain ($X^2(3)=10.12$, $p<.02$), the experts do more bounding of uncertainty and less planning. In the METOC domain ($X^2(3)=17.07$, $p<.001$), experts make more use of domain knowledge and mental spatial transformations, but do less examining of additional data. Of the six common strategies that differed in use across domains, it is interesting that 1) four of them also had expert/novice differences, and 2) the direction of difference is for experts across domains to differ more from one another than do novices. In other words, with practice in a domain, experts come to learn which general problem solving strategies are particularly useful in that domain (Lemaire & Siegler, 1995; Schunn, McGregor, & Saner, 2006).
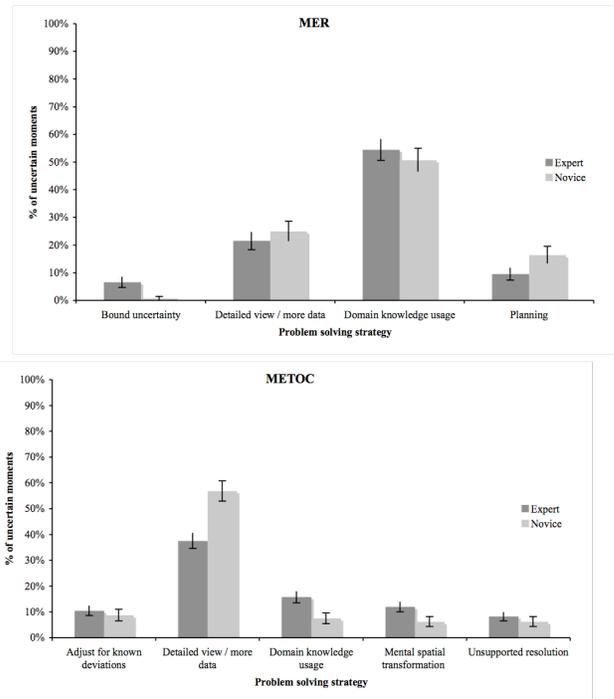

Figure 4: Expert/Novice relative use of uncertainty problem solving strategies in each domain.

While the common indicator strategies all had a clear spatial component, the common problem solving strategies varied significantly in terms of apparent degree of use of spatial representations. Of the five strategies that occurred often enough in the MER domain to assess degree of spatial representations (collapsing across expertise levels), only two strategies had spatial gestures significantly above the 18% of segments base rate overall during significant uncertain moments: mental spatial transformations (70%, $X^2(1)=16.12$, $p<.001$) and planning (33%, $X^2(1)=4.71$, $p<.05$). The other strategies co-occurred with spatial gestures somewhat below the overall base-rate: Bound Uncertainty (8%), More data (15%), and Domain knowledge usage (15%).

General Discussion

What is the nature of problem solving in complex domains? In this paper, we have focused on one element that contributes to the complexity of problem solving: informational uncertainty. Rather than focusing on traditional distinctions among cognitive processes (e.g., analogies, spatial reasoning, categorization, deductive reasoning, etc.), we take a strategies perspective (Siegler, 1996)—people solve tasks using a large variety of different strategies that are a rich interplay of more basic cognitive elements, and that choices in behavior and learning takes place at the level of these strategies.

Becoming Uncertain

We highlight four different strategies used by experts and novices in two different domains to identify significant

elements of uncertainty. The relative use these indicator strategies differed significantly across domains, likely reflecting the relative abundance of information provided—multiple sources likely inflates the use of inconsistency and novelty the occurrence of ambiguity.

The relative use of these indicator strategies did not differ significantly by expertise levels. However, the statistical power of the expertise contrasts for indicator strategies was not high. If the trends that were observed are to be believed, they are consistent with the notion that experts more strongly use the strategies that are particularly useful in a given domain.

Interestingly, the three most common strategies were significantly associated with spatial representations. The degree of association was above base-rates, but far from 100%. Measuring spatial representations using spatial gestures may result in an undercount of the level of spatial representation. However, level of spatial gestures for the indicator strategies was well below the highest level of spatial gesturing in one of the problem solving strategies, and thus is not likely just a conservative measure issue. More likely is that these indicators are not exclusively spatial, but rather a rich mixture of spatial and verbal reasoning. A related caveat is the spatial gestures may reflect communicative functions of gesture rather than purely a mirror of internal representation (McNeil, 1992). But this caveat simply reflects the general indirect nature of measures of internal representations.

Becoming Certain

How do problem solvers resolve informational uncertainty? Our analyses suggest that in complex real world settings, problem solvers use a rich array of strategies rather than any particular basic cognitive tool, like induction, deduction, analogy, categorization, or spatial reasoning. The same set of strategies applies across very different domains, although, again, the distribution of strategy use varies significantly by domain. While the number of variables that differ across the two domains observed here are much too great to convincingly argue for any particular factors underlying the differences in strategy base-rates by domain, we provide one possible explanation: the relative density of additional information to explore immediately. When the information is dense, one can spend considerable time exploring it for converging evidence, or use heuristics that focus on particular sources or average (mentally) across sources. When information is less dense, one must make guess based on domain knowledge or plan for additional data collection. Interestingly, the domain differences were exaggerated among experts (versus novices), suggesting a functional rather than coincidental difference by domain.

Some of our previous research in these domains heavily emphasized the role of spatial reasoning in resolving uncertainty. Indeed, we found that spatial reasoning goes up immediately following even minor moments of uncertainty (Schunn et al, 2007), especially in experts (Trafton et al., 2006). The current, more complete analysis of uncertainty

resolution problem solving suggests a more nuanced picture: spatial reasoning is certainly an element, but it appears to be just one tool among many, and that several of these other strategies appear not to be especially spatial in nature. But not all of the observed strategies occurred often enough to determine their spatial/verbal status, and future work should examine those strategies in detail.

Finally, the current work has examined just two different domains. Future work should examine a broader array of domains to understand what general features of domains influence the choice of uncertainty indicator and uncertainty problem solving strategies.

References

Christensen, B. T., & Schunn, C. D. (2007). The relationship of analogical distance to analogical function and pre-inventive structure: The case of engineering design. *Memory & Cognition 35*(1), 29-38.

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

Hutchins, E. (1995). *Cognition in the wild.* Cambridge, MA: MIT Press.

Kirsch, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18*(5), 513-549.

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General, 124*, 83-97.

McNeil, D. (1992). *Hand and Mind: What Gestures Reveal About Thought.* University of Chicago Press.

Newell, A., & Simon, H. A. (1972*). Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Schunn, C. D., McGregor, M., Saner, L. D. (2006). Expertise in ill-defined problem solving domains as effective strategy use. *Memory & Cognition, 33*(8), 1377-1387.

Schunn, C. D., Saner, L. D., Kirschenbaum, S. K., Trafton, J. G., & Littleton, E. B. (2007). Complex visual data analysis, uncertainty, and representation. In M. C. Lovett & P. Shah (Eds.), *Thinking with Data.* Mahwah, NJ: Erlbaum.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Trafton, J. G., Trickett, S. B., Stitzlein, C. A., Saner, L. D., Schunn, C. D., & Kirschenbaum, S. S. (2006). The relationship between spatial transformations and iconic gestures. *Spatial Cognition & Computation, 6*(1), 1-29.