THE DYNAMICS OF DIAGNOSING:
Virtuous and Vicious Cycles in the Operating Room

J. Bradley Morrison
Brandeis University
International Business School
bmorriso@brandeis.edu

Jenny W. Rudolph
Harvard Medical School
Center for Medical Simulation
jwrudolph@partners.org

John S. Carroll
Massachusetts Institute of Technology
Sloan School of Management
jcarroll@mit.edu

Summer 2008

The authors Jason Davis, Mark Healy, Zur Shapira, Karl Weick, the Academy of Management Review reviewers and editors, especially Olav Sorenson and Ron Adner, for their many helpful comments. Thanks also to the Boston College Organization Studies Alumni writing group (Erica Foldy, Pacey Foster, Danna Greenberg, Tammy MacLean, Peter Rivard, and Steve Taylor).

THE DYNAMICS OF DIAGNOSING:

Virtuous and Vicious Cycles in the Operating Room

Abstract

We develop a system dynamics model of diagnostic problem solving drawing on observation of doctors handling a medical emergency. The model links interpretation and choice, usually separated in the sensemaking and decision making literatures. Three insights emerge: (1) diagnostic problem solving includes acting, interpreting, and cultivating diagnoses; (2) dynamic feedback among these processes opens and closes windows of adaptive problem solving; and (3) reinforcing feedback processes, usually considered dysfunctional, are essential for adaptive problem solving. We discuss implications for improving theory and diagnostic problem solving in practice.

INTRODUCTION

An anesthesiologist is called to take over anesthesia in an operating room where a 29-year-old woman urgently needs an appendectomy. Soon the anesthesiologist notices the monitor is indicating that the patient's blood oxygen levels are falling below desired levels. The scenario presents a common but serious problem in anesthesia: difficulty with the process of ventilating, that is, breathing for the patient using a mechanical bellows. A variety of diagnoses for the ventilation problem are plausible such as an asthma attack, a collapsed lung, or insufficient paralyzing agent, but contradictory evidence is present for each, except one: the patient has exhaled some mucous into the tube, partially blocking it. The problem is not uncommon. Anesthesiologists have been acquainted with this problem in their training, but presentations of the problem can vary. Some air can get through the tube, but not enough for the patient to survive. Treatments addressing diagnoses other than the mucous plug in the breathing tube will not result in any sustained improvement in the patient's status. With a slowly dwindling level of oxygen in her blood, the patient can have uneven heartbeat and even go into cardiac arrest if the problem is not rectified. The cues the doctor has available include clinical signs and symptoms of the patient; new cues are generated by pursuing standard operating procedures for treating and diagnosing the patient.

Rudolph conducted an in-depth observational study of 39 doctors who faced this operating room crisis in a high-fidelity simulated operating room and found that only 9 correctly diagnosed the problem (Rudolph, 2003; Rudolph & Raemer, 2004). The other doctors fell prey to one of three dynamically distinct failure modes. The purpose of this paper is to examine the dynamics of how those 39 doctors handled the demands of dynamic problem solving in this time-pressured scenario with a (simulated) patient's life hanging in the balance. We set out to develop a dynamic model of the elements and feedback processes that generate variation in diagnostic problem solving, and our modeling has led to insights with implications for the more general domain of dynamic problem solving

The rest of the article is organized as follows: We first present the empirical findings that motivated modeling (Rudolph, 2003; Rudolph & Raemer, 2004). We then describe our iterative method of theory and model development, and review and synthesize the relevant research literatures in the exposition of our model. This exposition explains the key theoretical constructs of acting, interpreting and cultivating new diagnoses and the dynamic interactions among these constructs. Next, we discuss simulation results that provide insights on how adaptive and dysfunctional problem solving arises. We end with a discussion of the simulation results, highlighting the mechanisms that produce variation in dynamic problem solving, and their implications for problem solving theory and practice.

DYNAMIC PROBLEM SOLVING: AN EXAMPLE

The starting point for our theorizing is Rudolph's study 39 advanced anesthesia residents facing the simulated[1] operating room crisis described above (Rudolph, 2003; Rudolph et al., 2004). Rudolph observed problem solving by tracking doctors' concurrent verbal statements regarding their diagnoses, treatments and diagnostic tests and through post-hoc video review with participants. Rudolph found that the doctors fell into four modes

[1] We use the term "simulation" in two ways in this paper. The first use refers to the source data for Rudolph's study of clinical problem solving. These data were provided by a full-field, high-fidelity simulation (i.e. the research participant is in a fully equipped and staffed Operating Room (OR) with a computer controlled mannequin patient). The second use of the term refers to the computer-based simulation we conducted to analyze the system dynamics model.

of problem solving: stalled, fixated, vagabonding and adaptive (see Table 1 adapted from Rudolph, 2003). The doctors labeled *stalled* problem solvers had difficulty generating any diagnoses around which to organize action (on average just 1.5) and pursued few or no treatments and tests (1 on average). In contrast, those in the *fixated* mode quickly established a plausible, but erroneous, diagnosis to which they stuck despite countervailing cues (Table 1 shows that they considered their favorite diagnosis on average 10 times, double any other mode). Rather than pursuing multiple steps of a treatment algorithm to rule out diagnoses, they repeated the same step or did not advance through the treatment algorithm action steps. Although previous studies of fixation error (also known as premature closure or tunnel vision) generally conclude that broadening the range of alternatives considered is the needed antidote to fixation (Gaba, 1989; Johnson, Hassenbrock, Duran, & Moller, 1982), Rudolph's findings indicated that broadening could also be a problem: the data included a third mode labeled diagnostic *vagabonding*[2] in which doctors generated a wide range of plausible diagnoses and jumped from one to another without utilizing multiple action steps of the treatment algorithms (1.5 on average) for addressing and ruling out these diagnoses. Finally, the *adaptive* sensemaking mode, which looks very much like canonical models of effective clinical reasoning (Elstein et al., 1978), was characterized by generation of one or more plausible diagnoses and exploitation of multiple steps of known treatment algorithms. This process allowed those in the adaptive mode to rule out some diagnoses, take effective action and, unlike any other problem solving mode, resolve the breathing problem.

---

[2] This follows work of Dietrich Dörner, who identified a similar phenomenon among public officials attempting to identify effective strategies for public policy (Dörner, D. 1997. ***The Logic of Failure: Recognizing and avoiding error in complex situations***. New York: Perseus..

Table 1: Summary of Source Data

	Problem Solving Mode				
Variable	**Stalled**	**Fixated**	**Vagabonds**	**Adaptive**	Test of difference
N	**2**	**11**	**17**	**9**	–
Subjects who resolved the airway problem	**0**	**0**	**0**	**7**	$\chi 2$ (3) = 28.4***
Different Treatment Steps for a Diagnosis	**1.0 (0.0)**	**2.0 (1.1)**	**1.5 (0.5)**	**3.6 (0.7)**	$F(3,35) = 17.0$***
Considerations of Favorite Diagnosis	**3.0 (0.0)**	**10.0 (5.7)**	**5.4 (2.3)**	**5.9 (2.2)**	$F(3,35) = 5.0$**
Number of Different Diagnoses Considered	**1.5 (0.7)**	**3.8 (1.7)**	**6.1 (1.3)**	**5.0 (1.4)**	$F(3,35) = 9.1$***

Note -- means are given with standard deviation in parentheses. ** p < .01; *** *p* < .001

METHODS: SIMULATION FOR THEORY DEVELOPMENT

We used an iterative, grounded theory approach to develop our model using both data and theory as inputs using methods now well-articulated in the system dynamics community (Black, Carlile, & Repenning, 2004; Perlow et al., 2002; Repenning & Sterman, 2002; Rudolph et al., 2002; Sastry, 1997). Central constructs and relationships in our model emerged from constant comparison among our source data, related theories, and the emerging model as is characteristic of the grounded theory approach Strauss & Corbin, 1994; Glaser & Strauss 1967). More recently Davis, Eisenhardt, and Bingham (2007) described a seven-step "roadmap," for developing theoretical insights through simulation. Our approach, summarized in Table 2, overlaps greatly but differs in a few elements that are distinctive to building grounded models.

Table 2: Steps for Developing Grounded Theory Using Simulation Methods

Step	Activities
Frame problem	▪ Review source data ▪ Specify key dynamic behavior patterns of interest ▪ Determine a theoretically intriguing research question* ▪ Identify theories that address research question but where data are hard to obtain*
Conceptualize model	▪ Choose a simulation approach suited to dynamic processes* ▪ Choose model boundary ▪ Identify key constructs and assumptions ▪ Describe theoretical logic of causality ▪ Iterate among emerging model, source data, and relevant theory ▪ Draw causal loop diagram of feedback structure
Translate into mathematical model	▪ Operationalize theoretical constructs* ▪ Specify assumptions* ▪ Build system of equations that mirrors theoretical logic* (i.e., formulate and link equations for all variables) ▪ Use standard system dynamics formulations wherever possible
Simulate and analyze	▪ Conduct full set of model robustness checks, including extreme conditions tests* ▪ Replicate dynamic patterns in source data ▪ Design and conduct simulation experiments
Translate insights to written word	▪ Select simulation output for exposition ▪ Clearly explain how model structure causes observed behavior ▪ Interpret simulation results for theory and practice

* Items marked with an asterisk mirror steps or activities described in Davis, Eisenhardt, and Bingham's "Roadmap for Developing Theory Using Simulation Methods" (2007: 482).

With Rudolph's (2003, 2004) taxonomy of four diagnostic problem solving modes as a starting point, we followed the logic of grounded theory building (Strauss & Corbin, 1994), starting by identifying the constructs and relationships found in narratives of the problem solving modes and translating then into the language of stocks, flows, and feedback loops (e.g., Forrester, 1961; Sterman, 2000). We experimented with various causal loop diagrams, eventually converging on a rough draft that included the processes of acting, interpreting, and cultivating new diagnoses that ended up in the final model. In our second step, we compared the constructs and relationships in our diagrams with the sensemaking and decision making literatures. Third, we revised and streamlined our causal loop diagram, converging on just three central processes. Fourth, we translated the links and loops of this diagram into a formal mathematical model. The model's robustness and its ease of interpretation were enhanced by using standard system dynamics formulations (familiar fragments of model structure that occur frequently) wherever possible (Sterman, 2000). Finally, we simulated and analyzed the behavior of the model and used these analyses to elaborate a theory explicitly mapping the role of both balancing and reinforcing processes in effective and ineffective problem solving. While grounded in previous work, the model also provides new insights.

CONCEPTUAL MODEL

Overview

Our modeling process converged on the idea that dynamic problem solving involves three central tasks linked by feedback: 1) Problem solvers take actions and thus make information available for their interpretation, 2) They interpret the flow of information around them to continually reassess the plausibility of their diagnoses, and 3) They cultivate alternative diagnoses when current diagnoses are not satisfactory.

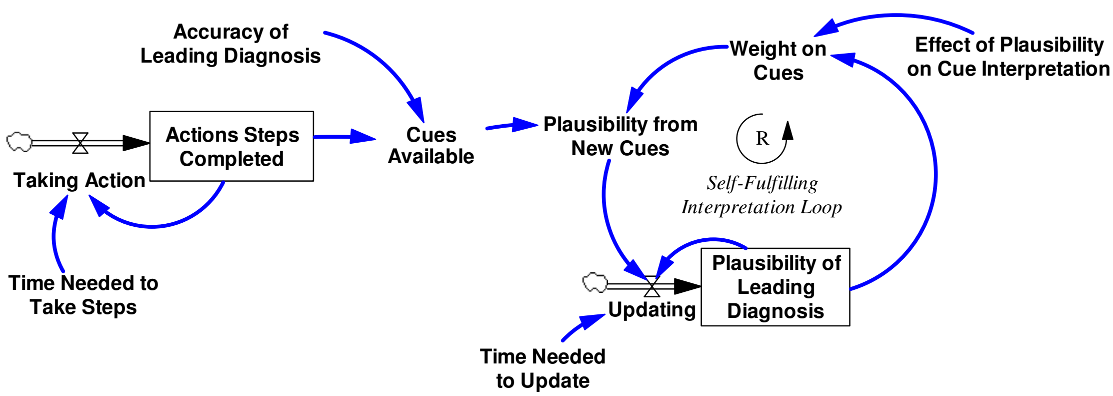
In our motivating example, the doctor expects the patient to breathe normally but instead observes and seeks to address a serious problem with the patient's ventilation. To succeed, the problem solver (an executive, a fire commander, etc.) must construct an organizing story about what is wrong (Weick et al., 2005). In our clinical example, the doctor scans the clinical signs and symptoms, patient history, and timing of the problem, and a plausible story develops in her mind; this takes the form of a diagnosis (Elstein, 2001; Elstein, Shulman, & Sprafka, 1978; Klein, Phillips, Rall, & Peluso, 2006; Rudolph et al., 2004). We now present the key constructs in our model of problem solving -- acting, interpreting, and cultivating new diagnoses -- and show how they are linked.

Dynamic problem solving requires taking action

Catalyzed by a surprise, interruption, or deviation from expectations and guided by an initial organizing diagnosis, the problem solver launches into action. In the domain of time-pressured acute medical care, conducting tests and providing treatments often involves following a standardized algorithm, a set of steps that combines therapeutic treatment with diagnostic tests (Cook & Woods, 1994; Elstein et al., 1978). This is a rule-based behavior that requires executing a known procedure (Rasmussen, Pejtersen, & Goodstein, 1994). Many other professional domains from nuclear power plant operations to computer chip manufacturing quality control have standard diagnostic or operating procedures to address problems (Carroll et al., 2002; Cyert & March, 1963; Edmondson, 2002; Gersick & Hackman, 1990; Levitt & March, 1988; Repenning et al., 2002; Winter, 1971). By moving through the steps of an algorithm or standard operating procedure, problem solvers generate cues that become available for them to notice, bracket, filter, and interpret. Having advanced the steps further, the problem solver has access to a larger pool of cues for making meaning in an ambiguous situation.

Thus, one important feature of acting is that the consequences accumulate. We model this progress as the stock *Action Steps Completed* that is increased by *Taking Steps,* as shown in Figure 1. Accomplishing action steps takes time (delays), sets the problem solver on a particular course of action (inertia), and yields results that remain available for interpretation (memory). The rate of *Taking Steps* depends on how many of the action steps remain to be executed and on the *Time Needed to Take Steps*, a parameter that represents the time needed for mental organizing to execute a step, physical rearranging to prepare for the step, executing the step, awaiting a response from the system, and noticing the results as cues in the stream of ongoing experience.

Figure 1
Feedback structure of acting and interpreting



As Plausibility of Leading Diagnosis *increases, the* Weight on Cues *decreases, so all else equal the* Plausibility from New Cues *continues to grow, leading to* Updating *that increases* Plausibility of Leading Diagnosis *still further. The process forms a reinforcing feedback loop, labeled R, the Self-fulfilling Interpretation Loop.*

A second important feature of action is that taking action steps makes *Cues Available* (see Fig. 1) for interpretation. For example, as the doctor accomplishes steps in a clinical treatment algorithm, she generates new diagnostic information that she can then consider in the context of her diagnosis. Analogously, as an organization carries out a new strategic initiative, executives can observe changes in the competitive environment that become inputs to their assessment of the situation and the merits of their strategies.

To model the stream of action-generated cues, we assume the information made available for interpretation is either confirming or disconfirming depending on which leading diagnosis is under consideration. We define a variable called *Accuracy of the Leading Diagnosis*, such that when the problem solver's leading diagnosis is correct, action-generated cues confirm the diagnosis and when it is incorrect, they are disconfirming. This feedback regarding environmental cues mimics the problem solving situation in a range of domains such as manufacturing defect elimination or computer debugging, while the simplifying assumption of correct or incorrect diagnosis allows us to place the mechanisms of interpretation in stark relief to clarify how they operate.

Dynamic problem solving requires interpreting cues

Facing a complex and ambiguous situation where quick action is needed, a problem solver has to create meaning in order to act. Unlike decision making experiments in a laboratory where "meaning already exists and is waiting to be found," in these settings meaning "awaits construction that might not happen or might go awry…" (Weick, 1995a: 15). Generating a plausible story or explanation about ambiguous cues helps organize and launch action (Neisser, 1976; Snook, 2000; Weick et al., 2005). Studies of strategic action (Sutcliffe et al., 2003), enactment of organizational structures (Weick, 1995a; Weick et al., 2005), naturalistic medical problem solving (Elstein, 2001; Elstein et al., 1978; Johnson et al., 1982), tactical decision-making under real-world stress (Cannon-Bowers & Salas, 1998), problem detection (Klein et al., 2005; Mandler, 1982), and problem solving in other naturalistic environments (Carroll, Rudolph, Hatakenaka, Wiederhold, & Boldrini, 2000; Klein, Orasanu, Calderwood, & Zsambok, 1993; Zsambok & Klein, 1997) all indicate that a plausible explanation, diagnosis or "story" is the engine of problem solving. Problem solvers such as corporate executives, chess players, or fire fighters, for example, use an initial diagnosis or assessment of the situation to develop a plan of action and decide what further information is needed (Dreyfus, 1997; Klein, 1998; Sutcliffe et al., 2003):

> It seems practically impossible to reason without hypotheses whenever the data base is as complex as it typically is in clinical problems. People are invariably trying to make sense out of their experience as it unfolds and are always generating hypotheses to explain their observations (Elstein & Bordage, 1979, p. x).

Once an initial diagnosis is established, how plausible people consider their own explanations or diagnoses waxes or wanes as they make sense of new or existing cues (Koehler, 1991; Smith & Blankenship, 1991). It takes time for people to notice, bracket, and label new cues and then change their mental models (e.g., diagnoses) accordingly (Bandura, 1991; Bartunek, 1984; Kleinmuntz, 1985; Marcus & Nichols, 1999; Roberts, 1990).

To capture this process of perceived plausibility increasing or decreasing with the interpretation of cues, we define *Plausibility of the Leading Diagnosis* as a stock variable shown in Figure 1. This stock depicts the problem solver's current subjective assessment of plausibility. By "leading" we mean the problem solver's favorite, or current most plausible, explanation. *Updating* is the process by which the current view of the *Plausibility of Leading Diagnosis* (the stock) is adjusted to equal the *Plausibility from New Cues*. *Updating* is an ongoing process of incorporating interpretations based on new information with current beliefs. Since changes in perceived plausibility require time, the rate of *Updating* is influenced by a parameter we call *Time Needed to Update*. *Plausibility from New Cues* describes interpretations of information generated by acting, which in turn depends in part on *Cues Available*

The links between cues and perceived plausibility

Making meaning from a stream of *Cues Available* through noticing, bracketing, filtering and labeling occurs while the problem solver holds a belief in the leading diagnosis, which may influence these interpretive processes. New assessments of plausibility (*Plausibility from New Cues*) depend not only on *Cues Available* but also on how open the problem solver is to external cues, which we label *Weight on Cues*. Studies of confirmation bias find that once an explanation is set, people prefer supporting to disconfirming information and this effect is stronger when cues are presented serially, as in our model (Jonas, Schulz-Hardt, Frey, & Thelen, 2001). Studies of fixation find that as the plausibility of the current diagnosis rises, openness to external cues, especially ones that defy the current view, decreases (De Keyser & Woods, 1990; Johnson, Moen, & Thompson, 1988; Staw, 1976; Xiao & MacKenzie, 1995). In other words, *Weight on Cues* is a downward-sloping function of *Plausibility of Leading Diagnosis*.

However, prior research is surprisingly silent regarding the exact form of the relationship between plausibility and weight given to external cues. For bold problem solvers, a small increase in plausibility leads to a disproportionately large decrease in weight on cues. For cautious problem solvers, a small increase in plausibility will lead either to no or only a small decrease in the weight on external cues. We use a parameter labeled *Effect of Plausibility on Cue Interpretation* to model the variation, from boldness to caution, in how plausibility influences the weight on cues. Appendix 1 depicts our representations of this relationship for different values of *Effect of Plausibility on Cue Interpretation*.

The recursive interactions between the interpreting and updating processes form a feedback loop. If new cues arrive that increase the *Plausibility of Leading Diagnosis*, then the *Weight on Cues* decreases slightly, leading to a small increase in the *Plausibility from New Cues*, which in turn causes *Updating* to further increase the *Plausibility of Leading Diagnosis*, and the cycle continues. This interpretation process amplifies a change through a reinforcing feedback process, labeled with the "R" for Reinforcing and named the "Self-Fulfilling Interpretation Loop." In the absence of any offsetting influences, this loop pushes the plausibility of an early-generated diagnosis toward ever greater plausibility. If the loop is driving toward greater plausibility of an erroneous diagnosis, it will generate the well-known self-confirming pattern of fixation, in which an initially plausible diagnosis and the filtering of external cues recursively influence each other so that the problem solver sticks to the diagnosis despite discrepant cues. If the loop is driving toward greater plausibility of a correct diagnosis, this is salutary. As we demonstrate later, the interplay between this interpretation process and the processes of acting and cultivating alternative diagnoses gives rise to the distinctive modes of dynamic problem solving observed among the anesthesiologists.

Dynamic problem solving requires cultivating new diagnoses

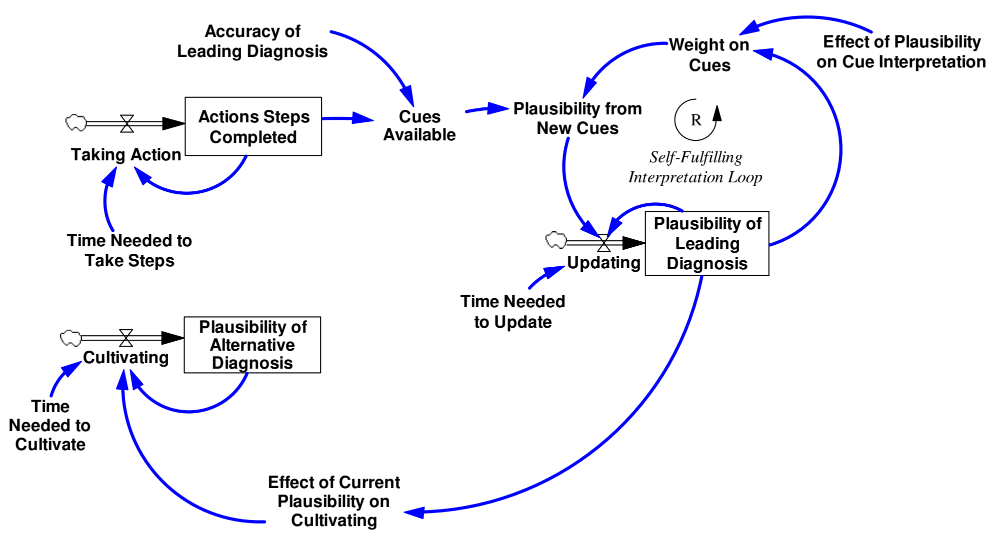
Problem solvers not only assess the plausibility of their leading diagnosis but also consider alternative diagnoses that are identified through search (Gupta, Smith, & Shalley, 2006; March, 1991), conversations (Weick et al., 2005), explanations (Hirt & Markman, 1995), or imagination (Amabile, 1982). This is a knowledge-based activity relying on expertise (Gonzales, Lerch, & Lebiere, 2003; Rasmussen et al., 1994). The process of inferring the most likely explanation for the available information is known as abduction (Josephson & Josephson, 1994; Peirce, 1958) and is complementary to deduction, or forecasting data that would be the consequences of a presumed hypothesis, and induction, or drawing conclusions from data. Together, abduction, deduction, and induction form a problem solving cycle of learning- or sensemaking-through-action.

The problem solving literature suggests a variety of ways for problem solvers to identify the most plausible hypothesis. At one extreme, all possible hypotheses are held simultaneously and updated through Bayesian inference as each new piece of information is received (Fischhoff et al., 1983). In well-structured clinical situations, for example, there are formal protocols that prioritize lists of diagnoses and clinical tests and treatments. Such decision rules are particularly useful for atypical problems or less experienced problem solvers (Gonzalez et al., 2003). The other extreme is represented by Klein's (1998) naturalistic decision making model that proposes that, for ill-structured problems representative of many real life situations, problem solvers consider only one hypothesis at a time, mentally simulate the implications of the hypothesis given the available information, and take action if the mental simulation confirms the plausibility of the hypothesis. Only if the mental simulation fails to confirm the hypothesis is a new diagnosis imagined and checked for validity. For ill-structured problems there are only bits and pieces of knowledge that may help a clever diagnostician find an obscure diagnosis or even invent a new diagnosis, and there may always be ambiguity about whether the diagnosis is correct. In the middle, Behavioral Decision Theory recognizes the constructive processes at work in decision behavior (Payne et al., 1992). And the problem-solving literatures include models of expmplar-based memory retrieval (Gonzalez et al. 2003, Logan, 1988) and the concept of a race between exemplars and general heuristics to produce a solution or response (Logan, 1988).

For our simulation model, we rejected both the extreme Bayesian rationality model and the serial diagnosis model. The behavioral decision making literature strongly challenges the realism of Bayesian updating, arguing that limited cognitive capacity makes such omniscience humanly impossible, even for well-structured problems (Fischhoff et al., 1983). Klein's model seems well-suited when the costs of information and the costs of erroneous diagnosis are very high, e.g., his fire commanders "test" their diagnosis by going into a building that may collapse around them. Klein's fire commanders are also highly expert and can "see" the answer using pattern matching with prior experience. On the other hand, if tests are easily run and erroneous diagnoses are easily replaced, and decision makers are not sufficiently expert to see a pattern (Elstein & Schwarz, 2002; Gonzales et al., 2003), problem solvers may hold multiple diagnoses in mind and seek further evidence to distinguish among them. Indeed, rather than training clinicians not to have hypotheses or to have all possible hypotheses, they are taught "differential diagnosis" in which they hold more than one (but not "all") diagnoses in mind and perform diagnostic tests that distinguish among the active diagnoses (Barondess & Carpenter, 1994). In Rudolph's study, more than half of the subjects are comparing simultaneous diagnoses. We therefore modeled a process that was psychologically reasonable and computationally simple, involving comparison of two potential diagnoses.

We assume that at any one point in time the problem solver has a preferred or leading diagnosis in mind and is seeking to validate or discredit that diagnosis (Elstein et al., 1978; Klein et al., 2006). Drawing on this research and Rudolph's data, we noted that doctors run tests (albeit with high variance in the quality) that will confirm a correct diagnosis or disconfirm an incorrect one. This process of running tests is not Bayesian updating in a mathematical sense, but rather some combination of logical rule-following (if I can put something down the breathing tube then it isn't blocked) and intuitively totaling up the supportive and countervailing evidence (Elstein et al., 1978; March, 1994). We further assume that there is always at least one other diagnosis that they are imagining (or taking from a well-learned list of candidates) and that the next best alternative among these others is gathering plausibility as the presenting problem (e.g., deteriorating oxygen status) remains unresolved. We label this process *Cultivating* and combine it with the processes of acting and interpreting in Figure 2. The *Plausibility of Alternative Diagnosis* is a stock that is increased by *Cultivating*, the pace of which is determined by the parameter *Time Needed to Cultivate.* We also show that the pace of *Cultivating* is reduced when the *Plausibility of Leading Diagnosis* is high by including a variable labeled *Effect of Current Plausibility on Cultivating*.

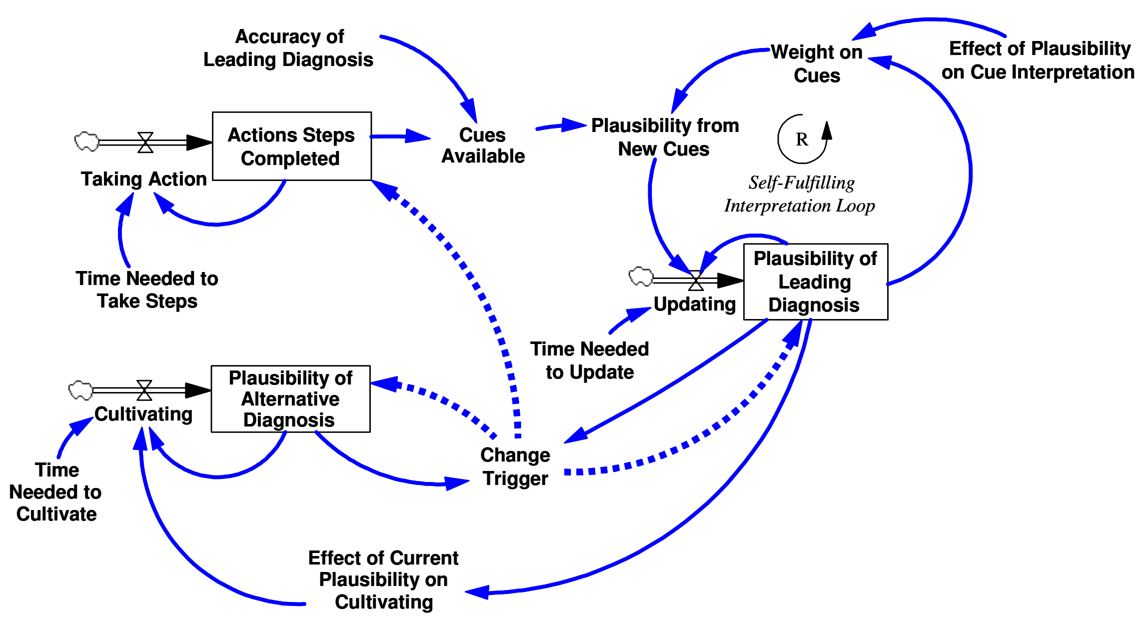
Figure 2
Core model structure showing the interaction of acting, interpreting, and cultivating alternatives



If the alternative diagnosis catches up to the leading diagnosis, we assume that the leading diagnosis is rejected. Three things happen when the leading diagnosis is rejected: 1) the alternative diagnosis becomes the new leading diagnosis, 2) the problem solver switches to action steps appropriate for the new leading diagnosis (so the *Action Steps Completed* starts over) and 3) yet another diagnosis becomes the second place alternative diagnosis (so *Plausibility of Alternative Diagnosis* starts over). We use dotted lines in Figure 3 to signal these changes when the leading diagnosis changes.

In summary, this model shows three problem solving processes: acting, interpreting, and cultivating new diagnoses. Together they bridge the gap between the sensemaking and decision making literatures by showing how meaning making and choice evolve and interact. We also incorporate both balancing and reinforcing processes.

Figure 3
Changes to reset the model when the alternative diagnosis becomes the leading diagnosis



We translated the causal structure represented in Figure 3 into a formal mathematical model so we could use simulation to pursue our theory development process. We drew upon the large body of standard formulations used to specify system dynamics models (Sterman, 2000) and filled in the gaps where necessary. We initially set parameters to reasonable values in the context of our motivating clinical example and refined them as we compared simulation output to the patterns in the clinical example data. We also conducted extensive sensitivity analysis to test for model robustness under extreme conditions and to explore the range of model behavior. Complete documentation of the model equations appears in Appendix 2.

SIMULATING THE DYNAMICS OF PROBLEM SOLVING

We begin with a set of experiments[3] that show how the interplay of acting, interpreting, and cultivating new diagnoses produces the four modes of diagnostic problem solving observed in Rudolph (2004). For clarity of exposition, we chose a scenario that controls for the effects of random search by assuming that all problem solvers generate alternative diagnoses in the same sequence and is consistent with the modal sequence in Rudolph's data. The simulations we present here are all based on a scenario in which the first, second, and third diagnoses considered are incorrect, the fourth is correct, and the fifth and all others after that are incorrect. Simulation analyses not shown here replicate the main results for scenarios in which the correct diagnosis enters earlier or later than fourth,

---

[3] "Experiment" is commonly used in the modeling community to refer to manipulations of the model parameters Carley, K. 2001. Computational approaches to sociological theorizing. In J. Turner (Ed.), ***Handbook of Sociological Theory***: 69-84. New York: Kluwer Academic/Plenum.

in which there are two correct diagnoses, and even in which the poor diagnoses have a modest degree of correctness.[4]

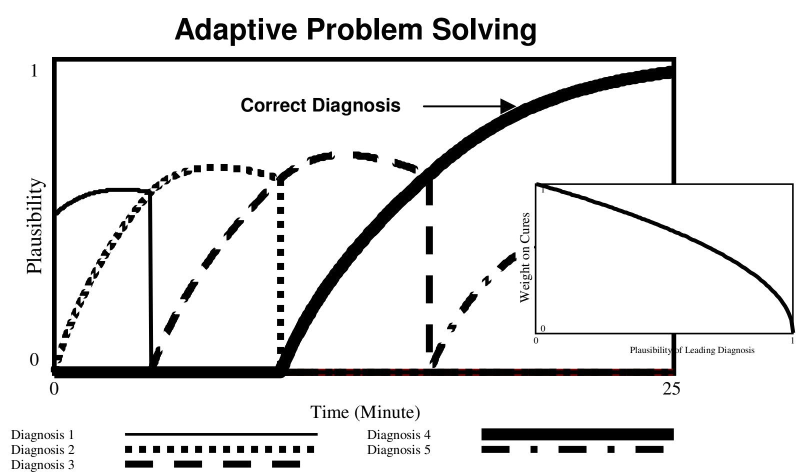
Four modes of dynamic problem solving

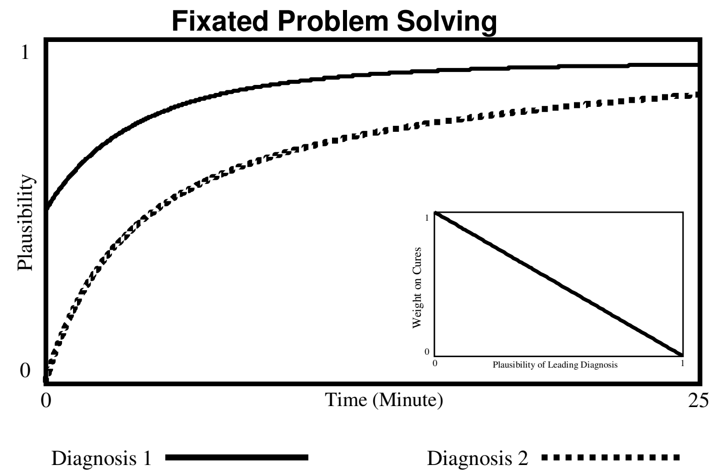
To highlight differences among the four problem solving modes we display in Figure 4 the behavior over time of the *Plausibility of the Leading Diagnosis*. The top panel is an illustration of the adaptive mode. The problem solver's sense of the plausibility of the first diagnosis begins at its initial value of 0.5 (out of 1.0), and three things begin to occur simultaneously. First, the problem solver begins taking action steps associated with the first diagnosis, increasing the stock of *Action Steps Completed*, which results in more *Cues Available*. Second, armed with some confidence in her diagnosis, the problem solver's interpretations begin to increase plausibility, and the *Weight on Cues* begins to fall slowly as the Self-Fulfilling Interpretation Loop acts to reinforce the leading diagnosis. In the first few moments, the diagnostic algorithm has not progressed much, so the limited cues have little effect on plausibility. After a short time, the accumulated cues (which are "objectively" disconfirming information because the first diagnosis is incorrect) begin to show their effect on plausibility, and we see a slow decline in the *Plausibility of the Leading Diagnosis*. Third, the plausibility of an alternative diagnosis builds as the cultivating process unfolds in the face of cues unfavorable to the leading diagnosis. Eventually, plausibility of the alternative overtakes the *Plausibility of the Leading Diagnosis*. At this moment the first diagnosis is rejected and the second diagnosis becomes the leading diagnosis. The problem solver begins pursuing the algorithm associated with the new leading diagnosis. The pattern repeats for the second

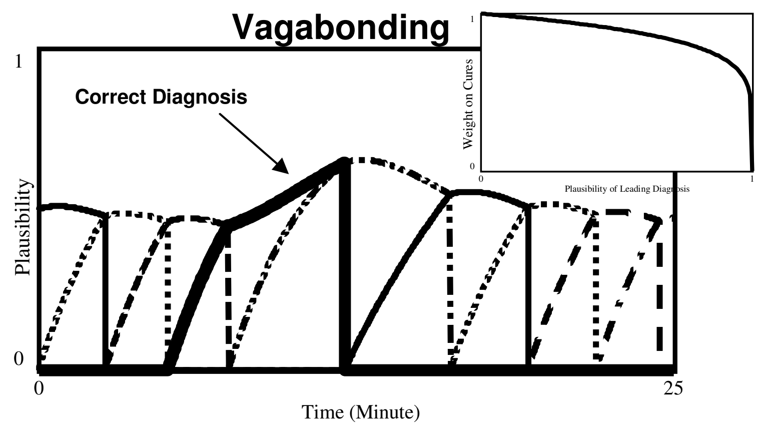
[4] A summary of several hundred additional simulations demonstrating these results is available from the first author on request.

and third diagnosis: plausibility increases for a short while, disconfirming cues accumulate and begin to cause a reduction in plausibility, and an alternative diagnosis gains favor and eventually overtakes the leading diagnosis. When the problem solver begins to consider diagnosis number four, the correct one, plausibility begins to grow as before. However, the new *Cues Available* now offer confirmation and are interpreted to build even more *Plausibility of Leading Diagnosis*. Moreover, the Self-Fulfilling Interpretation Loop reinforces the increases in plausibility, reducing the *Weight on Cues* thus boosting plausibility still further. The diagnostician pursues the action steps to completion and converges on a steady state choice of the correct diagnosis.

Figure 4: Modes of Dynamic Problem Solving







From top to bottom: the adaptive, fixated, and vagabonding modes of problem solving. Simulation conditions are identical except for the strength of the self-fulfilling interpretation feedback loop, as determined by the relationship between Plausibility of the Leading Diagnosis and the Weight on Cues, shown in the inserts.

This mode of behavior shows two important features of problem solving dynamics. First, the consideration of each diagnosis enjoys a honeymoon period during the time it takes for an alternative diagnosis to emerge as a viable contender. In the adaptive mode, this temporal interplay between the leading and alternative diagnoses is “well-balanced” in that the honeymoon period is long enough for the problem solver both to take action and to interpret the results for both the incorrect and correct diagnoses she considers. Second, there is a dynamic interchange in the roles of acting and interpreting because the cues available accumulate slowly relative to the ongoing process of interpreting experience.

Plausibility increases at first because updating driven by the confirmation-biased interpretation process occurs quickly relative to the accumulation of available cues. Meanwhile, the problem solver continues to take action and generate more cues. As the disconfirming evidence mounts, it eventually overcomes the effects of the self-fulfilling interpretation: Plausibility reaches a peak and then begins to decline.

The middle panel of Figure 4 shows simulation results that replicate the fixating mode. The difference between this experiment and the one in the top panel is only that the *Effect of Plausibility on Cue Interpretation,* depicted in the inset, is stronger.[5] Plausibility starts to rise as before at first, but the lower *Weight on Cues* in this scenario allows the self-fulfilling process to gain momentum. The problem solver acts, interpreting cues and creating meaning that supports the current diagnosis, and the *Weight on Cues* falls even more. The Self-Fulfilling Interpretation Loop reinforces the current diagnosis, and because the loop is so strong, the first diagnosis is always preferred. The diagnostician does not move on to any other diagnosis. The strong reinforcing effects result in a pattern of problem solving in which the problem solver is completely confident in the incorrect diagnosis. Self-fulfilling interpretations discount some disconfirming evidence, so the current diagnosis locks in prematurely, squeezing out the cultivation of alternatives, and the problem solver never has a chance to find the correct diagnosis.

The bottom panel in Figure 4 shows simulation results that replicate the vagabonding mode. In this experiment, the *Effect of Plausibility on Cue Interpretation* is weaker[6]. The first three diagnoses are rejected, but more quickly than in the adaptive case,

---

[5] Specifically, *Effect of Plausibility on Cue Interpretation* = 1

[6] Specifically, *Effect of Plausibility on Cue Interpretation* = 0.15

implying that the problem solver does not take as many action steps, consistent with field data showing that diagnostic vagabonds generated diagnoses but performed few or no steps of the treatment/test algorithm (1.5 steps on average). When the fourth diagnosis, the correct one, enters as the leading diagnosis, plausibility increases, but not as rapidly as in the adaptive case. The problem solver places a higher *Weight on Cues* (due to a weaker *Effect of Plausibility on Cue Interpretation*), but the cues to confirm the diagnosis accumulate somewhat slowly because they must be made available by advancing through action steps. Meanwhile, an alternative diagnosis gains favor and eventually overtakes the correct diagnosis, and the problem solver also rejects the correct diagnosis number four. Once this diagnosis is rejected, the problem solver continues identifying alternatives, choosing them as the leading diagnosis, and rejecting them in favor of the next emerging alternative.

The stylized problem solver in the vagabonding experiment is quite capable of cultivating alternatives and attending to cues, but lacking more confident beliefs about the plausibility of a diagnosis, she does not hold onto it long enough to adequately advance forward with action steps. The error in this mode is the premature rejection of the correct diagnosis number four. The result is vagabonding, a pattern of diagnostic problem solving in which the problem solver jumps from one plausible diagnosis to the next without treating the patient. The dynamic interplay among acting, interpreting, and cultivating alternatives is out of balance: the problem solver yearns for clarifying information (interpreting) but the pace of generating new cues associated with the leading diagnosis (acting) is too slow relative to the pace of cultivating alternatives. The problem solver gets stuck in a mode of generating new alternatives but not discovering enough about them to reach an effective conclusion. This mode fails because the effect of plausibility on interpretation is so weak that even the correct diagnosis is rejected.

The model can also generate a mode in which the problem solver is stalled, unable to move forward to take any action steps (for example, when both taking action steps and cultivating diagnoses are extremely slow processes). Rudolph's analysis classified only 2 out of 39 doctors as stalled. Both exhibited behaviors of advancing treatment steps little or not at all and establishing working diagnoses very slowly, consistent with this example. However, with so few examples and so little action to learn from, we omit this mode from subsequent analysis.

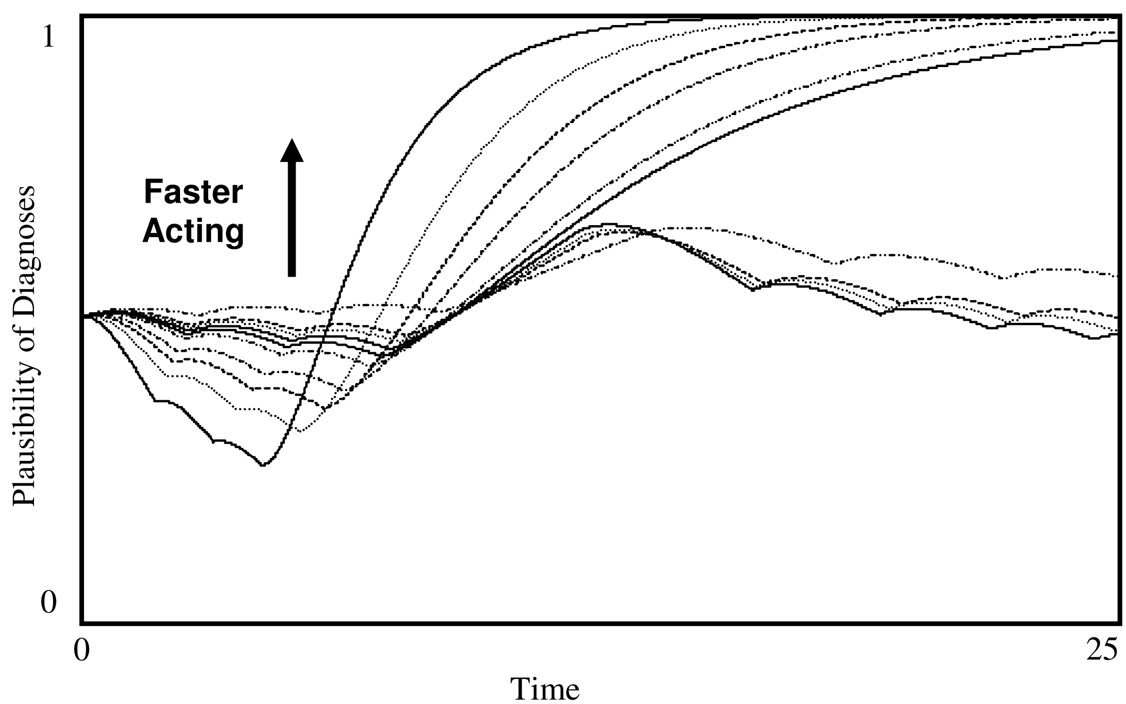
The interplay of acting, interpreting, and cultivating diagnoses: Sensitivity analysis

To examine the critical dynamic interactions among acting, interpreting, and cultivating diagnoses more closely, we conducted experiments in which we varied the pace of these processes. The first set of simulations holds all parameters the same as in the vagabonding case of Figure 4 except that we vary the pace of acting[7]. We discovered that the eleven simulation runs we generated, shown in Figure 5, separate into two distinct patterns. The set corresponding to faster acting is adaptive: the plausibility of diagnosis number four climbs smoothly toward one. The other set, based on slower acting, displays vagabonding: diagnosis number four is rejected and new alternatives continue overtaking the lead. This experiment points to two important results about the system's dynamics. First, different rates of taking action generate qualitatively different dynamics. The problem solver converges on the correct diagnosis for fast rates of action, but rejects the correct diagnosis when the rate is too slow. A bias for action (taking

[7] By setting the *Time Needed to Take Steps* to values ranging from very fast (1 minute) to very slow (16 minutes)

action steps faster) offsets the effects of less self-fulfilling interpretation and protects the problem solver from being swept into vagabonding mode. Second, small differences in the rate of acting can mean the difference between adaptive sensemaking and vagabonding. This result raises the question as to just what pace of taking action is needed to escape from the perils of vagabonding.

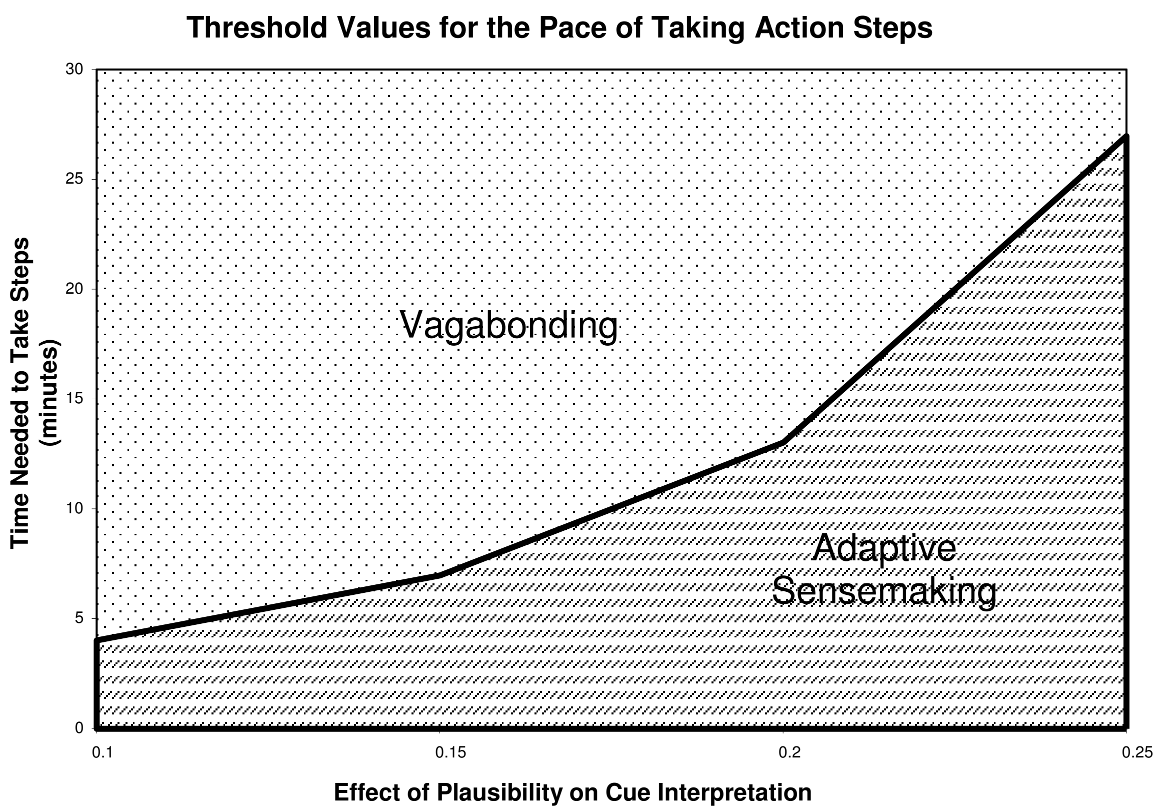
Figure 5
Sensitivity analysis showing system behavior for various rates of taking action steps



To shed light on this question, we conducted an extensive set of experiments to test the relationship among the pace of acting, the pace of cultivating alternatives, and the strength of the effect of plausibility on interpretation. We performed many sets of simulations like those in Figure 5 for various values of the *Effect of Plausibility on Cue Interpretation* and found the threshold pace of *Time Needed to Take Steps* needed to achieve adaptive sensemaking for the given combination of the other two parameters.

The results, displayed in Figure 6, show that for weaker *Effects of Plausibility on Cue Interpretation*, faster action taking is needed for adaptive sensemaking. A weak interpretation effect describes a problem solver who wants more cues, so the pace of acting must be faster in order to lead to adaptive sensemaking. When the appetite for cues is high (weak effect), slow action induces vagabonding. Conversely, a modest degree of confidence in the leading diagnosis contributes to the robustness of the sensemaking process by thwarting the lurking threat of vagabonding.

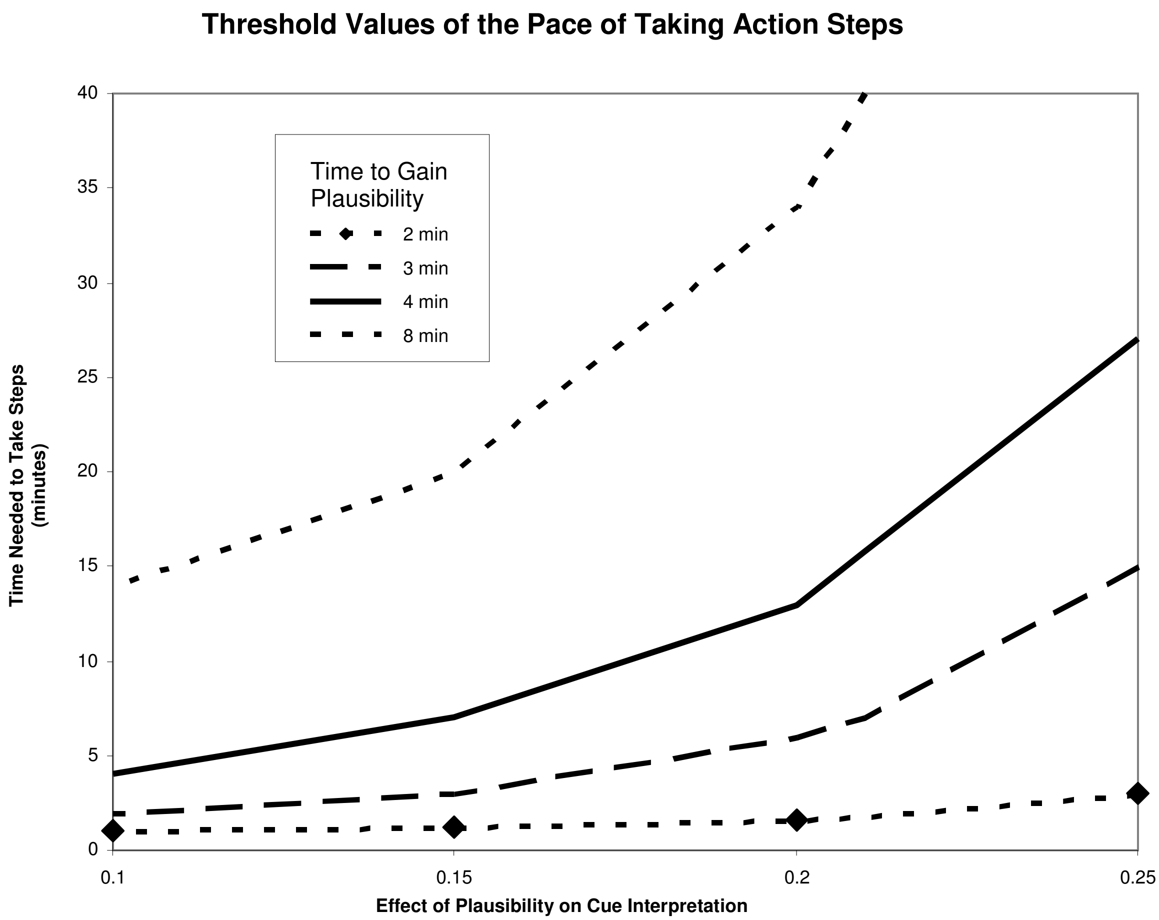
Figure 6



To characterize the dynamic interplay among the processes of acting, interpreting, and cultivating alternatives fully, we repeated the analysis for several values of the *Time*

Needed to Cultivate. The resulting family of curves (Figure 7) shows how the threshold pace of taking action depends on both the *Effect of Plausibility on Cue Interpretation* and the *Time Needed to Cultivate*. As the strength of the interpretation effect increases, the threshold pace of acting needed for adaptive problem solving gets slower. When the pace of cultivating alternatives is very fast, the risk of vagabonding is quite high and not mitigated much by stronger interpretation effects. Very rapid action is still needed. For a slower pace of cultivating alternatives, small increases in the strength of the interpretation effect yield larger improvements in the robustness of the problem solving process. Slower paces of action are still adequate to achieve adaptive problem solving.

Figure 7



DISCUSSION

Our simulation model generates the four modes of diagnostic problem solving identified in Rudolph's study. Stalling occurs when problem solvers are not confident enough in their diagnosis to launch action. Fixation occurs when plausibility in the leading diagnosis builds too quickly and new information is interpreted as confirmation, preventing alternative diagnoses from being considered sufficiently. Vagabonding occurs when the leading diagnosis does not gain plausibility rapidly enough and is rejected too quickly, even when correct. Adaptive problem solving exhibits a balance between giving the leading diagnosis enough plausibility to sustain concerted action, but not too much to prevent alternatives from emerging. In our various model tests and experiments, we did not observe additional modes of problem solving.

Our modeling and grounded theory development process contributes three new insights to understanding diagnostic problem solving. First, we clarify and delineate three processes linked by feedback that are central to dynamic problem solving: acting, interpreting, and cultivating diagnoses. Second, while current theories of sensemaking and choice look separately at the above processes, our modeling shows how they are inextricably linked with tradeoffs and interactions between processes. Third, we highlight the fact that reinforcing feedback, usually seen as the driver of maladaptive patterns such as fixation, can actually play a beneficial role in adaptive problem solving, depending on its intensity. In this discussion, we also comment on our grounded theory approach to using simulation models for theory development and close the discussion with thoughts on practical implications for professional problem solving and future research directions.

Simulation Models and Theory Development

The theory development literature often emphasizes the field's lack of clarity on how to generate theory, while emphasizing the importance of theory building (Lave & March, 1993; Sutton & Staw, 1995; Weick, 1995). The simulation theory development literature places heavy emphasis on the validity of models, but provides little guidance on the specific methods for developing grounded models from data, especially where feedback dynamics are involved (Davis, Eisenhardt, & Bingham, 2007; Sterman, 2000). Our approach facilitates the elaboration and extension of previous theories in a highly transparent way, relying on inductive and iterative theory building and elaboration rather than deducing the model from general principles (Black et al., 2004; Davis et al., 2007; Rudolph et al., 2002). Insights about the domain of dynamic problem solving emerged through a constant cycle of comparison and refinement that involved invention, explication of assumptions and relationships through model articulation, and experimentation guided and constrained by existing theory and data (see Table 3). The model served as a "boundary object," a shared space that allowed the authors to clarify and surface their emerging theoretical insights, translate these into diagrams and equations, and explore their implications through simulation (Carlile, 2002; Star et al., 1989).

Table 3: Insights from the Modeling Process

Steps	Insights
Frame problem	• Diagnostic problem solving is a dynamic process in which candidate diagnoses emerge and then gain or lose favor over the course of the problem solving time horizon. • Alternative diagnoses compete to become the problem solver's focus of attention and action. • Action and cognition occur in parallel and interact.
Conceptualize model	• Dynamic problem solving comprises three interacting processes of acting, interpreting, and cultivating alternatives. • The feedback structure of dynamic problem solving includes both adaptive feedback processes, as balancing loops, and one potentially destabilizing process, as a reinforcing loop. • The fact that information can be gained only through action introduces delays which contribute to important dynamics.
Translate into mathematical model	• The self-fulfilling interpretation effect can vary in strength. • Plausibility both launches action and also fosters interpretations consistent with current beliefs. • The leading diagnosis is a special focus of action.
Simulate and analyze	• Weak effects of plausibility on interpretation can explain diagnostic vagabonding. • Acting, interpreting and cultivating alternatives are three levers problem solvers can use to compensate for weaknesses in the other processes and convert failure modes to adaptive problem solving. • Plausibility has a role in both launching and sustaining action.
Translate insights to written word	• A key boundary condition is the need to take action to discover information. • Our theory has practical implications, for example, in the training of doctors. • Our learning benefited from the dialogue and discourse of our group modeling process.

This mode of theory development has certain characteristics which are becoming familiar in the simulation community. First, the process helps identify implicit assumptions. For example, this effort highlighted the fact that previous models of sensemaking and adaptive decision making assume balancing feedback and ignore the role of reinforcing processes. Second, it highlights new implications of existing theory. For example, we found that the links among acting, interpreting, and cultivating new diagnoses can exacerbate dysfunction in each other or can compensate for each other. Third, the constraints of mathematical modeling enforce a rigorous articulation of internally consistent theory, which paradoxically generates new substantive insights. For example, our model demonstrated that there are windows of opportunity for adaptive problem solving that can open and close. Finally, although formal mathematical models are criticized for losing the nuances of thick description, many narrative theories and associated analyses rely on relatively static, simple relationships and interactions. Formal modeling pushed us to find functional forms that offer an enriched articulation of relationships. For example, modeling the process of self-fulfilling interpretation or confirmation bias led to the discovery of a nuanced non-linear relationship between perceived plausibility and weight on cues.

Improving Problem Solving in Practice

Whereas researchers and practitioners have previously identified a range of strategies to reduce the risks of falling prey to fixation, our work suggests that vagabonding is another failure mode that should be avoided. Success in training managers and clinicians to avoid fixation may even increase the tendency to vagabonding. Situational factors most likely to lead to vagabonding include high urgency, high stakes, delays in gathering information, and many plausible alternatives. When these factors are coupled with slower or less confident interpretation processes, as might be expected with novel problems or relatively inexperienced problem solvers, the risk of vagabonding will be greater. Potential strategies to avoid vagabonding are to slow down the pace of cultivating alternatives, take action more confidently, and hold leading diagnoses more confidently to allow more cues to surface. These three strategies interact and support each other, so improvement in one dimension can compensate for a shortfall along another.

Our simulation results point to the importance of constraints on the core processes of acting, interpreting, and cultivating alternatives. The pace of action might be limited by resource constraints, technological factors, or physical factors that could be adjusted once they are understood. Problem solvers may be more able to avoid self-fulfilling interpretation effects if trained how to be mindful of or "by stand" their own diagnostic frames rather than mistaking them for reality (Kegan, 1994; Langer, 1989; Torbert, 1991). The pace of generating alternatives could be strengthened by reducing competing demands for attention and providing knowledge and experience relevant to the problem at hand. Just as a virtuoso musician will learn to play over a range of loudness, the versatile problem solver will develop the ability to adjust the pace of acting, interpreting, and cultivating alternatives to match the needs of the situation (e.g., Gonzalez et al., 2003). However, such expertise presumably develops over considerable time and exposure to a variety of situations.

Limitations and Future Research

Given the interacting constructs and relationships in the model, there is more than one way to generate the problem solving behaviors observed by Rudolph (2003). For example, fixation can be produced by increasing the confirmation bias, but also by slowing down the generation of new hypotheses. We experimented with some model variations, using as criteria the fit to the source data, the psychological realism in comparison with the problem solving literature, and the parsimony of varying as few model elements as possible to reproduce all four modes. Our argument is not that our model is an exact representation of dynamic problem solving, but rather that the variations in model behavior have been sufficiently realistic and informative to provide insights about theory.

However, to continue the dialogue between modeling and empirical data, more of each is needed. We could experiment with other versions of our model and continue to examine various bits and pieces of theory in order to simulate the source data in a behaviorally and psychologically realistic way, but there is fundamentally too little data and too many ways to model problem solving in this complex situation. The modeling exercise does provide a direction for identifying data that would be most useful for sorting out possibilities. We know from the model that speed and timing of the three major processes matter a great deal, and we also know from the literature that the level of knowledge of the doctors (residents are intermediate) is likely to shift how they process data (ability to do differential diagnosis or even "see" the correct diagnosis more readily). By articulating mechanisms in even a stylized way, we have identified areas for further research, for example, how plausibility changes with new data, how hypotheses are generated, how actions are chosen, and how stress (e.g., time pressure) has impact. Additional focused studies examining a variety of mediating variables (e.g., timing, hypothesis generation, plausibility, discounting) with appropriate methods (e.g., process tracing techniques) could be used to identify these processes more directly.

CONCLUSION

This paper developed a theory about the role of plausibility in dynamic problem solving grounded in existing theory and empirical data. Simulation of the system dynamics model of the theory showed how underlying structures and relationships produce the various problem solving modes: stalling, fixating, vagabonding and adapting. The formal modeling process helped extend existing problem solving theory in three ways: 1) It has clarified core dynamic elements of inertia and change in acting, interpreting, and generating new diagnoses. 2) It has taken theories of problem solving that assert interactions among acting, interpreting, and cultivating diagnoses and represented the interactions explicitly (e.g., how the pace of generating new cues influences assessments of plausibility and the need for cultivating alternate diagnoses), allowing examination of windows of opportunity for adaptive sensemaking. 3) Most importantly, it has generated new insights suggesting that problem solving theory must include both balancing and reinforcing processes. Specifically, the formal modeling process has allowed us to demonstrate the benefits and nuances of self-fulfilling interpretation; some fixation-like activity is needed for adaptive diagnostic sensemaking (but too much can cause problems). Through modeling we have also demonstrated that the specific form of the relationship between faith in the plausibility of the current diagnosis and openness to new cues is more complex than previous theories of sensemaking and fixation have appreciated.

We are far from a complete and comprehensive theory of dynamic problem solving. However, the modeling process and grounded theory approach has identified areas in which additional data would help distinguish among alternative models. These are the most likely growth points for exploring this fertile ground for future research.

REFERENCES

Amabile, T. M. 1982. Social psychology of creativity: A consensual assessment technique. ***Journal of Personality and Social Psychology***, 43(5): 997-1013.

Arthur, W. B. 1990. Positive feedbacks in the economy. ***Scientific American***, Feb.: 92-99.

Bandura, A. 1991. Social cognitive theory of self-regulation. ***Organization Behavior and Human Decision Processes***, 50: 248-287.

Barondess, J., & Carpenter, C. (Eds.). 1994. ***Differential Diagnosis***. Philadelphia Lea & Febiger Lea & Febiger.

Bartunek, J. M. 1984. Changing interpretive schemes and organizational restructuring: The example of a religious order. ***Administrative Science Quarterly***, 29: 355-372.

Black, L., Carlile, P., & Repenning, N. P. 2004. A dynamic theory of expertise and occupational boundaries in new technology implementation: Building on Barley's study of CT scanning ***Administrative Science Quarterly***, 49(4): 572-607.

Cannon-Bowers, J. A., & Salas, E. 1998. ***Making Decisions under Stress: Implications for Individual and Team Training***. Washington, DC: American Psychological Association.

Carley, K. 2001. Computational approaches to sociological theorizing. In J. Turner (Ed.), ***Handbook of Sociological Theory***: 69-84. New York: Kluwer Academic/Plenum.

Carroll, J. S., Rudolph, J. S., Hatakenaka, S., Wiederhold, T. L., & Boldrini, M. 2000. Incident investigation team diagnoses and organizational decisions at four nuclear power plants. In G. Klein (Ed.), ***Naturalistic Decision Making***. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Carroll, J. S., Rudolph, J. W., & Hatakenaka, S. 2002. Learning from experience in high-hazard industries. ***Research in Organizational Behavior***, 24: 87-137.

Cook, R. I., & Woods, D. 1994. Operating at the sharp end: The complexity of human error. In B. S. Bogner (Ed.), ***Human Error in Medicine***: 255-310. Hillsdale, NJ: Lawrence Earlbaum Associates.

Cyert, R. M., & March, J. G. 1963. ***A Behavioral Theory of the Firm***. Englewood Cliffs, NJ: Prentice Hall.

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. 2007. Developing theory through simulation methods. ***Academy of Management Review***, 32(2): 480-499.

De Keyser, V., & Woods, D. D. 1990. Fixation errors: Failures to revise situation assessment in dynamic and risky systems. In A. G. Colombo, & A. Saiz de Bustamante (Eds.), ***Systems Reliability Assessment***: 231-251. Amsterdam: Kluwer.

Dörner, D. 1997. ***The Logic of Failure: Recognizing and avoiding error in complex situations***. New York: Perseus.

Dreyfus, H. L. 1997. Intuitive, deliberative, and calculative models of expert performance. In C. E. Zsambok, & G. Klein (Eds.), ***Naturalistic Decision Making***: 17-28. Mahway, NJ: Lawrence Erlbaum Associates.

Edmondson, A. 2002. Disrupted routines: Team learning and new technology implementation in hospitals. ***Administrative Science Quarterly***, 46(4): 685-716.

Eisenhardt, K. M. 1989. Making fast strategic decisions in high-velocity environments. ***Academy of Management Journal***, 32(3): 543-576.

Elstein, A. S. 2001. Naturalistic decision making and clinical judgment. ***Journal of Behavioral Decision Making***, 14(5): 363-365.

Elstein, A. S., & Schwarz, A. 2002. Evidence base of clinical diagnosis: Clinical problem solving and diagnostic decision making: selective review of the cognitive literature. 324(7339): 729-732.

Elstein, A. S., Shulman, L. S., & Sprafka, S. A. 1978. ***Medical Problem Solving: An Analysis of Clinical Reasoning***. Cambridge, MA: Harvard University Press.

Fischhoff, B., & Beyth-Marom, R. 1983. Hypothesis evaluation from a bayesian perspective. ***Psychological Review***, 90(3): 239-260.

Fleming, L., & Sorenson, O. 2004. Science as a map in technological search. ***Strategic Management Journal***, 25: 909-928.

Gaba, D. M. 1989. Human error in anesthetic mishaps. ***International Anesthesiology Clinics***, 27(3): 137-147.

Gersick, C. J. G., & Hackman, J. R. 1990. Habitual routines in task-performing groups. ***Organizational Behavior and Human Decision Processes***, 47: 65-97.

Gonzales, C., Lerch, J. F., & Lebiere, C. 2003. Instance-based learning in dynamic decision making. ***Cognitive Science***, 27: 591-635.

Gupta, A. K., Smith, K. G., & Shalley, C. E. 2006. The interplay between exploration and exploitation. ***Academy of Management Journal***, 49(4): 683-706.

Hirt, E. R., & Markman, K. 1995. Multiple explanation: a consider-an-alternative method for debiasing judgments. ***Journal of Personality and Social Psycholgoy***, 69(6): 1069-1086.

Hogarth, R. M. 1981. Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. ***Psychological Bulletin***, 90: 197-217.

Johnson, P. E., Hassenbrock, F., Duran, A. S., & Moller, J. H. 1982. Multimethod study of clinical judgment. ***Organizational behavior and human performance***, 30: 201-230.

Johnson, P. E., Moen, J. B., & Thompson, W. B. 1988. Garden path errors in diagnostic reasoning. In L. Bolc, & M. J. Coombs (Eds.), ***Expert System Applications***. Berlin: Springer-Verlag.

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. 2001. Confirmation bias in sequential information search after preliminary decisions: an expansion of dissonance theoretical research on selective exposure to information. ***Journal of Personality and Social Psycholgoy***, 80(4): 557-571.

Josephson, J. R., & Josephson, S. G. (Eds.). 1994. ***Abductive Inference: Computation, Philosophy, Technology***. New York Cambridge University Press.

Klein, G. 1998. ***Sources of Power***. Cambridge, MA: MIT Press.

Klein, G., Phillips, J. K., Rall, E., & Peluso, D. A. 2006. A Data/Frame Theory of Sensemaking. In R. Hoffman (Ed.), ***Expertise Out of Context***. Mahway, NJ: Erlbaum.

Klein, G., Pliske, R., Crandall, B., & Woods, D. 2005. Problem detection. ***Cognition, Technology and Work***, 7(1): 14-28.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. 1993. ***Decision Making in Action***. Norwood, NJ USA: Ablex Publishing.

Kleinmuntz, D. N. 1985. Cognitive heuristics and feedback in a dynamic decision environment. ***Management Science***, 31(6): 680-702.

Koehler, D. J. 1991. Explanation, imagination, and confidence in judgment. ***Psychological Bulletin***, 110(3): 499-519.

Langley, A., Mintzberg, H., Pitcher, P., Posada, E., & Saint-Macary, J. 1995. Opening up decision making: the view from the black stool. ***Organization Science***, 6(3): 260-279.

Lave, C. A., & March, J. G. 1993. ***An introduction to models in the social sciences***. Lanham, MD: University Press of America.

Levitt, B., & March, J. G. 1988. Organizational learning. ***Annual Review of Sociology***, 14: 319-340.

Louis, M. R., & Sutton, R. I. 1991. Switching cognitive gears: From habits of mind to active thinking. ***Human Relations***, 44(1): 55-76.

Mandler, G. 1982. Stress and thought processes. In S. Goldberger, & S. Breznitz (Eds.), ***Handbook of Stress***: 88-164. New York: Free Press.

Mandler, G. 1984. ***Mind and Body***. New York: W. W. Nortorn & Company.

March, J. G. 1991. Exploration and exploitation in organizational learning. ***Organization Science***, 2(1): 71-87.

March, J. G. 1994. ***A Primer on Decision Making: How Decisions Happen***. New York: Free Press.

Marcus, A. A., & Nichols, M. 1999. On the edge: Heeding the warning of unusual events. ***Organization Science***, 10(4): 482-499.

Neisser, U. 1976. ***Cognition and Reality: Principles and Implications of Cognitive Psychology***. San Francisco: W.H. Freeman and Company.

Payne, J. W., Bettman, J. R., & Johnson, E. J. 1992. Behavioral Decision Research: A Constructive Processing Perspective. 43(1): 87-131.

Peirce, C. S. 1958. ***Collected Papers***. Cambridge, MA: Harvard University Press.

Perlow, L., Okhuysen, G., & Repenning, N. P. 2002. The speed trap: Exploring the relationship between decision making and temporal context. ***Academy of Management Journal***, 5: 931-955.

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. 1994. ***Cognitive Systems Engineering***. J. Wiley and Sons: New York.

Repenning, N. P., & Sterman, J. D. 2002. Capability traps and self-confirming attribution errors in the dynamics of process improvement. ***Administrative Science Quarterly***, 47: 265 - 295.

Roberts, K. H. 1990. Some characteristics of one type of high reliability organization. ***Organization Science***, 1(2): 160-176.

Rudolph, J. W. 2003. ***Into the big muddy and out again: Error persistence and crisis management in the operating room***. Boston College, Chestnut Hill, MA.

Rudolph, J. W., & Raemer, D. B. 2004. Diagnostic problem solving during simulated crises in the OR. ***Anesthesia and Analgesia***, 98(5S): S34.

Rudolph, J. W., & Repenning, N. P. 2002. Disaster dynamics: Understanding the role of quantity in organizational collapse. ***Administrative Science Quarterly***, 47: 1-30.

Sastry, M. A. 1997. Problems and paradoxes in a model of punctuated organizational change. ***Administrative Science Quarterly***, 42: 237-275.

Simon, H. 1957. ***Models of man: Social and rational***. New York: Wiley.

Smith, S. M., & Blankenship, S. E. 1991. Incubation and the persistence of fixation in problem solving. ***American Journal of Psychology***, 104(1): 61-87.

Snook, S. A. 2000. ***Friendly Fire: The Accidental Shootdown of US Black Hawks Over Northern Iraq***. Princeton, NJ USA: Princeton University Press.

Starbuck, W. H., Greve, A., & Hedberg, B. L. T. 1978. Responding to crises. ***Journal of Business Administration***, 9(2): 111-137.

Staw, B. M. 1976. Knee-deep in the Big Muddy: A study of escalating commitment to a chosen course of action. ***Organizational Behavior and Human Performance***, 16: 27-44.

Sterman, J. 2000. ***Business Dynamics***. Chicago: Irwin-McGraw Hill.

Sterman, J. D. 1994. Learning in and about complex systems. ***System Dynamics Review***, 10(2-3): 291-330.

Strauss, A., & Corbin, J. 1994. Grounded theory methodology: An overview. In N. K. Denzin, & Y. S. Lincoln (Eds.), ***Handbook of qualitative research***: 273-285. Thousand Oaks: Sage.

Sutcliffe, K., & Weber, K. 2003. The high cost of accurate knowledge. ***Harvard Business Review***, May: 74-87.

Sutton, R. I., & Staw, B. M. 1995. What theory is not. ***Administrative Science Quarterly***, 40: 371-384.

Torbert, W. R., & Associates. 2004. ***Action Inquiry: The Secret of Timely and Transforming Leadership***. San Francisco: Berret-Koehler Publishers.

Weick, K. E. 1993b. The collapse of sensemaking in organizations: The Mann Gulch disaster. ***Administrative Science Quarterly***, 38: 628-652.

Weick, K. E. 1995. What theory is not theorizing is ***Administrative Science Quarterly***, 40(3): 385-390.

Weick, K. E. 1995a. ***Sensemaking in Organizations***. Thousand Oaks, CA: Sage.

Weick, K. E., Sutcliffe, K., & Obstfeld, D. 2005. Organizing and the process of sensemaking. ***Organization Science***, 16(4): 409-421.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. 1999. Organizing for high reliability: Processes of collective mindfulness. In R. I. Sutton, & B. M. Staw (Eds.), ***Research in Organizational Behavior***, Vol. 21: 81-123. Stamford, CT: JAI Press.

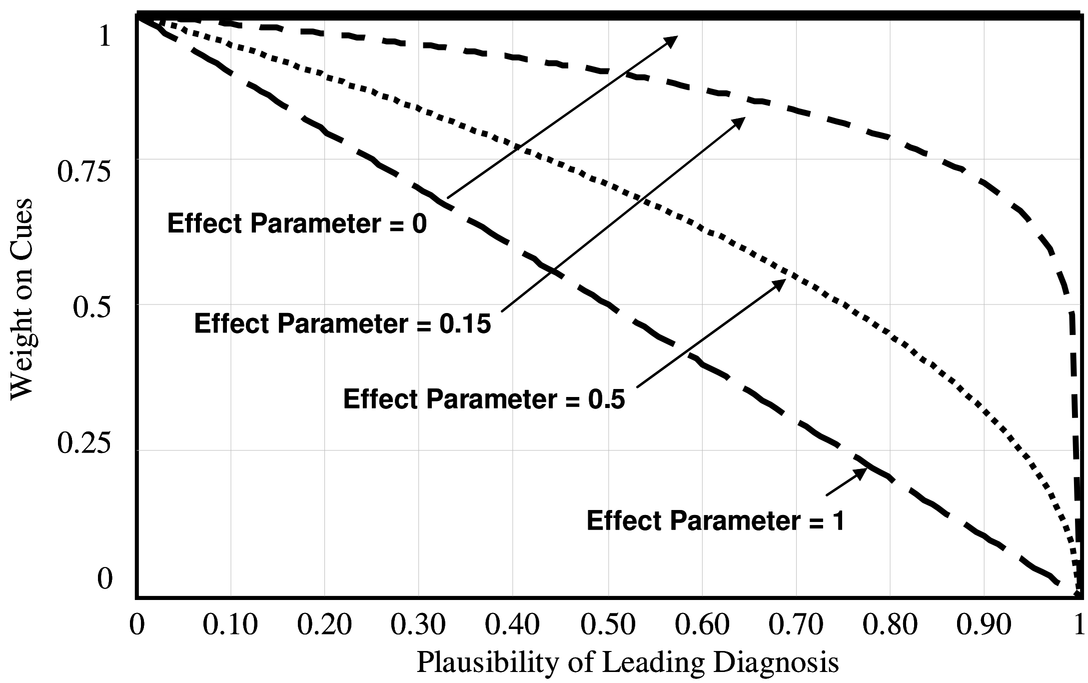
Winter, S. G. 1971. Satisficing, selection, and the innovating remnant. ***Quarterly Journal of Economics***, 85: 237-261.

Xiao, Y., & MacKenzie, C. F. 1995. ***Decision making in dynamic environments: Fixation errors and their causes***. Paper presented at the Human Factors and Ergonomics Society 39th Annual Meeting, Santa Monica, CA.

Zsambok, C. E., & Klein, G. 1997. ***Naturalistic Decision Making***. Mahway, NJ: Lawrence Erlbaum Associates.

APPENDIX 1

Figure A-1
Weight on Cues as function of Plausibility of Leading Diagnosis
for Various Settings of Effect of Plausibility on Cue Interpretation



When Plausibility of Leading Diagnosis *is at its extreme value of 0 (the problem solver perceives their leading diagnosis as completely implausible), the* Weight on Cues *is at its extreme value of 1 (the problem solver pays full attention to cues). Conversely, when* Plausibility *equals 1, the Weight on Cues is equal to zero (except when the* Effect of Plausibility on Cue Interpretation *is equal to zero, in which case no matter how plausible or implausible the problem solver deems his current diagnosis, he always gives full weight to cues).*

When the Effect of Plausibility on Cue Interpretation *is equal to .15, even large increases in the* Plausibility of Leading Diagnosis *do not much diminish the weight the problem solver places on cues. It is not until he is almost completely certain the diagnosis is plausible that his Weight on Cues diminishes. When the* Effect of Plausibility on Cue Interpretation *is .5, decreases in* Weight on Cues *are greater for a given increase in* Plausibility of Leading Diagnosis.*; When the* Effect *is equal to 1, a decrease in plausibility brings about a proportional decrease in* Weight on Cues.

Or, formally,

> Weight on Cues (t) = (1 - Plausibility of Leading Diagnosis (t)) ^ Effect of Plausibility on Cue Interpretation

where the exponent Effect of Plausibility on Cue Interpretation is a parameter chosen to represent possible individual and/or situational differences.

APPENDIX 2

Integral equations are written in this appendix using the following notation:
Stock= INTEG (Inflow-Outflow, Initial Value of Stock), where the INTEG function means the integral from time 0 to time *t* (the current time) of the inflow less the outflow plus the initial value of the stock. The model is simulated using Vensim DSS software, available from www.vensim.com.

Equations for the Acting Subsection (Figure 1):

Action Steps Completed= INTEG (Taking Action - Resetting Action Steps ,0)
Units: Dimensionless

Taking Action =(1- Action Steps Completed)/Time Needed to Take Steps
Units: Dimensionless/Minute

Time Needed to Take Steps =8
Units: Minute

Cues Available=(Starting Plausibility of Leading Diagnosis+ Action Steps Completed*(Accuracy of Leading Diagnosis- Starting Plausibility of Leading Diagnosis))
Units: Dimensionless

Accuracy of Leading Diagnosis=IF THEN ELSE(Current Diagnosis=True Diagnosis, 1, 0)
Units: Dimensionless

True Diagnosis=4
Units: Dimensionless

Equations for the Interpreting Subsection (Figures 2 and 3):

Plausibility of Leading Diagnosis= INTEG (Updating +Carry Over to Leading -Resetting Leading, Initial Plausibility)
Units: Dimensionless

Updating=(Plausibility from New Cues-Plausibility of Leading Diagnosis)/Time Needed to Update
Units: Dimensionless /Minute

Plausibility from New Cues=Cues Available*Weight on Cues+(1-Weight on Cues)
Units: Dimensionless

Time to Needed Update=2
Units: Minute

Weight on Cues=(1-Plausibility of Leading Diagnosis)^Effect of Plausibility on Cue Interpretation
Units: Dimensionless

Effect of Plausibility on Cue Interpretation=0.5
Units: Dimensionless

Equations for the Cultivating Alternatives Subsection (Figure 4)

Plausibility of Alternative Diagnosis= INTEG (Cultivating-Resetting Alternative,0)
Units: Dimensionless

Cultivating=Effect of Current Plausibility on Cultivating*(1-Plausibility of Alternative Diagnosis)/Time Needed to Cultivate
Units: Dimensionless/Minute

Effect of Plausibility on Alternative=min(1,2-2*Plausibility of Leading Diagnosis)
Units: Dimensionless

Time Needed to Cultivate=4
Units: Minute

Equations for Switching Diagnoses (Figure 5):

Change Trigger=IF THEN ELSE(Plausibility of Leading Diagnosis<Plausibility of Alternative Diagnosis, 1, 0)/TIME STEP
Units: Dimensionless/Minute

Resetting Action Steps =Action Steps Completed*Change Trigger
Units: Dimensionless/Minute

Resetting Leading=Plausibility of Leading Diagnosis*Change Trigger
Units: Dimensionless/Minute

Carry Over to Leading=Resetting Alternative
Units: Dimensionless/Minute

Resetting Alternative=Plausibility of Alternative Diagnosis*Change Trigger
Units: Dimensionless/Minute

Starting Plausibility of Leading Diagnosis = INTEG (New Plausibility-Resetting Starting Plausibility, Initial Plausibility)
Units: Dimensionless

New Plausibility=Resetting Alternative
Units: Dimensionless/Minute

Resetting Starting Plausibility=Change Trigger* Starting Plausibility of Leading Diagnosis
Units: Dimensionless/Minute

Initial Plausibility=0.5
Units: Dimensionless

Current Diagnosis= INTEG (Diagnosis Counter,1)
Units: Dimensionless

Diagnosis Counter=Change Trigger
Units: Dimensionless/Minute