

Gary Klein · Rebecca Pliske · Beth Crandall
David D. Woods

Problem detection

Received: 11 April 2004 / Accepted: 29 July 2004 / Published online: 22 December 2004
© Springer-Verlag London Limited 2004

Abstract Problem detection is the process by which people first become concerned that events may be taking an unexpected and undesirable direction that potentially requires action. Previous accounts [e.g., Cowan (Acad Manage Rev 11(4):763–776, 1986)] described problem detection as the accumulation of discrepancies until a threshold was reached. In reviewing incidents taken from a variety of natural settings, we found that discrepancy accumulation did not apply to the incidents we reviewed, because (a) cues to problems may be subtle and context-dependent, and (b) what counts as a discrepancy depends on the problem-solver's experience and the stance taken in interpreting the situation. In many cases, detecting a problem is equivalent to reconceptualizing the situation.

1 Introduction

Problem detection is critical for the effective management of complex, real-world situations. Problems must

G. Klein (✉) · B. Crandall
Klein Associates Inc.,
1750 Commerce Center Boulevard North,
Fairborn, OH 45324, USA
E-mail: gary@decisionmaking.com
Tel.: +1-937-8738166
Fax: +1-937-8738258
E-mail: beth@decisionmaking.com

R. Pliske
Dominican University, 7900 West Division Street,
River Forest, IL 60305, USA
E-mail: rpliske@email.dom.edu
Tel.: +1-708-5246658

D. D. Woods
Cognitive Systems Engineering Laboratory,
Ohio State University, 210 Baker Systems,
1971 Neil Avenue, Columbus, OH 43210, USA
E-mail: woods.2@osu.edu
Tel.: +1-614-2921700
Fax: +1-614-2927852

be recognized before actions can be taken to resolve them. The ability to detect problems at early stages can lead to more timely and effective interventions. Conversely, failures of early problem detection can result in accidents and performance breakdowns if action is not initiated until the situation has deteriorated to the point where recovery is impossible.

We rely on problem detection to provide us with an early warning under different types of conditions. In carrying out a plan, we need to be sensitive to potential problems as we construct the plan, and also as we execute the plan. While engaging in a routine activity such as driving, we need to notice disturbances that might signal a traffic jam or a hazardous condition. Even in a steady state condition, we need to be alert to possible dangers such as a tree branch that looks like it might fall on the roof of our house or a maintenance action that could affect the safe operation of a petrochemical plant.

Once people detect a problem, they can act in a variety of ways. They may seek more information, track the events more carefully, try to diagnose or identify the problem, raise the concern with other people (e.g., a nurse informing a physician that a baby is in trouble), to explain away the anomaly or take the initiative to cope with the problem by finding an action that would counter the trajectory of events; or they may accept that the situation has changed in fundamental ways, and revise the goals and plans.

2 What is problem detection?

Smith (1989) distinguished between *problem detection*, the initial factors that arouse concern, and *problem identification*, which results in the ability to specify the problem. Our research interest has been in the initial discovery that events are taking an unacceptable trajectory and may require action. For example, the initiating condition can be the unexpected appearance of a threat, or the non-appearance of a safeguard. We are not

concerned here with the attempt to identify the nature of the problem, because that shifts the focus from problem detection to the representation and diagnosis of a problem after it is detected. Sometimes, we can distinguish problem detection, problem identification and diagnosis as separate activities, but in many cases problem detection, problem definition and diagnosis occur together, and in other cases, diagnosis may not be needed.

Descriptions of problem solving in the psychological literature tend to pass over cognitive aspects of problem detection or to subsume problem detection as part of a general problem definition function (Anderson 1993; Davis 1973; Duncker 1945; Forbus and de Kleer 1993; Greeno and Simon 1988; Hayes 1981; Newell and Simon 1972; Polya 1957; Rubinstein 1975; Wertheimer 1959). The process of problem detection may seem like a straightforward triggering of the more complex cognitive functions of problem solving (e.g., problem identification, diagnosis, construction, and evaluation of one or more courses of action). In studies carried out in well-controlled settings, problem detection is often eliminated; participants are presented with the problem to solve, and do not have to discover it.

A further source of confusion is that researchers have used a variety of terms to refer to the initial processes involved in problem solving: problem detection, problem discovery, problem finding, anomaly detection (Woods et al. 1987), problem recognition (Cowan 1986; Schrenk 1969), crisis perception (Billings et al. 1980), and problem sensing (MacCrimmon 1973).

To illustrate the phenomenon of problem detection, we describe an incident studied by Crandall and Getchell-Reiter (1993, see Klein 2004 for a more complete account). In this first case, two nurses were appraising the same cues, but forming different judgments. The case is described from the viewpoint of the experienced nurse.

2.1 Case 1: An experienced versus an inexperienced nurse

When this incident took place, I was serving as an instructor for a new nurse. We had been working together for quite a while and she was nearing the end of her orientation; so, she was really doing primary care and I was in more of a supervisory position. Anyway, we were nearing the end of a shift and I walked by this particular isolette and the baby really caught my eye. The baby's color was off and its skin was mottled. It looked funny (belly slightly rounded). I looked at the chart and it indicated the baby's temperature was unstable. I also noticed that the baby had a heel stick for lab work several minutes ago and the stick was still bleeding. When I asked the orientee nurse how she thought the baby was doing, she said that he seemed kind of sleepy to her. I got the doctor immediately, told him we were "in big trouble" with

this baby. I said the baby's temperature was unstable, that its color was funny, it seemed lethargic, and it was bleeding from a heel stick. He reacted right away, put the baby on antibiotics and ordered cultures done. I was upset with the orientee that she had missed these cues, or that she had noticed them but not put them together. When we talked about it later I asked about the baby's temperature dropping. She had noticed it, but had responded by increasing the heat in the incubator. She had responded to the "surface" problem, instead of trying to figure out what might be causing the problem. The temperature had dropped each time over four readings and she had not realized the significance of the pattern at all.

This case illustrates how the underlying problem—the development of sepsis in the baby—was inferred from a set of different kinds of symptoms. However, the experienced nurse was not simply accumulating symptoms. Her first glance at the infant told her something might be wrong. Instead of walking past the infant, she began studying it in more detail, eventually looking at the temperature record and asking about the baby's condition. The identification of several indicators made the experienced nurse more concerned than if the baby was only showing a single symptom, but the mere accumulation of symptoms was not the basis for the nurse's reaction. She detected the problem from the very first glance, when she saw that the baby's color was off and its skin was mottled. The additional symptoms fit a mental model, an explanatory scheme, of how sepsis is manifested. The same symptoms were available to the instructor as to the orientee. However, the experience of the instructor allowed her to catch the pattern to the anomalies—the color being off, the mottled skin, and the shape of the belly. The instructor could also look at the data about the baby's temperature and see a trend that fit the pattern she was recognizing. To the new nurse, the falling temperatures meant that the baby was getting cold. To the experienced nurse, the same data placed in a configuration of cues meant that the baby was getting sick.

2.2 Case 2: Going for the feint

This incident occurred during a naval battle group exercise. The AEGIS cruiser expected a raid of about 40 aircraft and was not overly surprised when it was notified of six air contacts inbound at about 250 miles. Large raids often have the aircraft flying in smaller flights of six to eight aircraft. The commander was not overly suspicious seeing this small number. The assessment was that this was the lead flight, soon to be followed by the main raid.

Four aircraft were sent to intercept this lead element outside the 200-mile range. Additional aircraft were sent aloft to replace the ones sent to conduct the intercept, and to prepare for the follow-on raid. All six of the enemy aircraft were splashed at about 200 miles out.

Shortly thereafter, another raid was detected approaching from a different bearing. This turned out to be the main raid, with over 30 aircraft. The first raid was only a feint, meant as a distraction.

As the main raid was discovered late, the AEGIS cruiser was not able to intercept at a safe range. Since the AEGIS cruiser failed to maintain the outer-air battle, it deteriorated into an inner-air battle, requiring very challenging coordination among different elements assigned rapidly maneuvering targets.

Case 2 shows another incident of problem detection, but one that does not require expertise. The indicators are clear and obvious—the approaching aircraft of the main raid were easily picked up on radar. The problem was detected without much difficulty. The breakdown was in identifying the problem and understanding what was happening, not in noticing that an attack had begun.

We are not particularly interested in these types of cases because they are so straightforward. Incidents such as case 2 make it convenient to ignore problem detection in favor of the more difficult functions of identifying and diagnosing problems. Case 1 shows how problem detection can be very difficult and worth a more careful investigation.

3 Cowan's discrepancy accumulation model of problem detection

To date, the most comprehensive account of problem detection is provided by Cowan (1986), who presented a three-stage model of what he called the “problem recognition process.” It is based on earlier work showing that problem recognition is triggered by a discrepancy between the perceived existing state and models of what that state ought to be (Billings et al. 1980; Downs 1967). While we accept much of Cowan's account, particularly his description of factors that affect problem detection/recognition, evidence we will review suggests that his account is too limited. According to Cowan, the core of the problem recognition process is the accumulation of discrepancies between what is being observed and what is desired. These discrepancies accumulate until they pass some threshold and are noticed. The first stage of Cowan's model, which he labeled the gestation/latency stage, is where these discrepancies accumulate. During the second or categorization stage, the accumulated discrepancies are classified as “a problem” or “not a problem.”

Cowan's model was not based on empirical data, nor did it use analyses of actual problem detection events. In order to gain a better perspective on problem detection, we reviewed incident accounts gathered during cognitive task analysis interviews with experienced personnel drawn from a number of different fields. Our objective was to collect and review incidents, particularly difficult cases, in order to identify some of the major factors that affect the process of problem

identification. We relied on a naturalistic research approach as a means of formulating hypotheses about problem detection, rather than conducting a quantitative test of hypotheses.

4 Data collection and analysis methods

4.1 Re-analysis of incident accounts

Our primary source of data was a large set of critical incidents that had been accumulated in a number of our previous research efforts. We reviewed a large set of more than 1,000 incidents built up from multiple studies using the critical decision method (CDM) for cognitive task analysis (Hoffman et al. 1998; Klein et al. 1989). The CDM is an extension of Flanagan's critical incident method (Flanagan 1954). In addition to eliciting the critical incidents, the CDM is designed to probe these retrospective accounts of challenging events that typically require decision making and problem solving. The CDM is a semi-structured approach that first elicits a brief overview of the challenging incident, followed by a systematic account to develop a timeline of events. This is next elaborated to identify the key judgments and decisions, and to establish the information available for making these. Finally, hypothetical questions and other types of queries are used to further examine the cognitive processes involved in handling the incident.

An interview guide is prepared in advance of these data-collection sessions, but this guide is used only as a general framework. No attempt is made to ask each respondent the same questions, in the same order. Instead, follow-up questions are posed on the spot, depending on the previous responses. The interviews examine the types of information and data the participants recall using, rather than asking the participants why they made certain judgments or decisions.

As with any introspective method, the potential exists for memory distortions, and so the results of a CDM interview are not treated as accurate accounts of the incidents. Instead, they are treated as a source of hypotheses. CDM interviews are part of a naturalistic approach to studying cognition in field settings (Klein et al. 2003).

Incident accounts were reviewed from a variety of different CDM interview projects. Crandall and Calderwood (1989) conducted CDM interviews with 19 neonatal intensive care unit nurses, who carefully watch newborn infants for early signs of distress. Case 1 above was taken from this project. Pliske et al. (2004) report the results of a set of CDM interviews with 37 weather forecasters, who attempt to anticipate problematic weather conditions, particularly those that can affect air operations. Kaempf et al. (1996) used CDM interviews with US Navy Commanders and Tactical Action Officers to investigate early detection and reaction to threatening events such as the Vincennes incident that

resulted in a mistaken shoot-down of a commercial airliner in the Persian Gulf in 1988. Case 2 was taken from this project. Klinger and Militello (2002) used CDM interviews to elicit challenging incidents faced by Weapons Directors on board AWACS aircraft (Airborne Warning and Command System); most of these incidents came out of Operation Desert Storm, the 1991 war to free Kuwait. Klein et al. (1988) conducted CDM interviews with 26 experienced Fireground Commanders, eliciting 32 separate critical incidents. Klein and Hutton (1995) interviewed 11 scientists and engineers working for the Air Force Research Laboratory, to examine 15 incidents in which key problems were discovered and addressed. We also studied other observations of problem detection in actual or simulated cases during anomaly response in space shuttle mission control, process control rooms, anesthetic management during surgery, and aviation flight decks (Watts-Perotti and Woods 1997; Wood 1994).

In addition to these published accounts, we also reviewed unpublished project accounts involving en route Air Traffic Controllers and Navy Landing Signal Officers.

4.2 New critical incident interviews

We also conducted new CDM interviews that were specifically directed at examining the problem detection process. The new critical incidents came from two domains: wildland firefighting (five interviews) and minimally invasive surgery for removal of the gallbladder (three interviews).

Wildland firefighting was chosen because firefighters must constantly be vigilant about their own safety and the safety of their team. To this end they are trained to set up safety zones to which they can retreat if necessary. Safety zones are areas of burned out forest that no longer contain fuel to sustain the fire, or roads or other paved areas or clearings where fire cannot spread. The issue that raises questions about problem detection is when to retreat to a safety zone. The CDM interviews were conducted with commanders who had been recently in situations where they had needed to retreat to a safety zone.

We also conducted three interviews with surgeons experienced with the procedure of laparoscopic cholecystectomy. Laparoscopy is a term for a gallbladder removal procedure in which a small camera and long thin instruments are inserted through tiny incisions in the body. This is in contrast to an open procedure in which a large incision is made to admit the surgeon's hands and instruments. With a laparoscopic procedure, the surgeon can misrecognize anatomical structures and injure other structures near the gallbladder. Sometimes the structures (ducts and arteries) cannot be identified clearly. One option is to convert the procedure from laparoscopic to open to permit direct handling of the tissues and a direct binocular view of the anatomy.

Problem detection enters here as surgeons become concerned about difficulties in visualizing the anatomy and consider whether to convert to an open procedure.

Dominguez and her colleagues (Dominguez 1998; Dominguez et al. 2004) had investigated the judgment to convert to an open procedure during laparoscopic cholecystectomy by collecting think-aloud protocols with surgeons as they watched several videotapes of actual surgeries and discussed rules, cues, predictions, concerns, comfort level, metacognition, and perceptual expertise. We reviewed transcripts from this study and then conducted three interviews with surgeons as they viewed videotapes of this type of surgery.

Our intention in conducting these reviews and performing a few additional CDM interviews was to identify incidents that shed light on the problem detection process as it occurs in natural settings. Inasmuch as we were not initiating a data collection effort to obtain new cases, we were not concerned with data coding or frequency counts for different categories of responses. The review was intended simply to gain a qualitative understanding of the problem detection process, and to generate hypotheses. We were interested in accounts of either successful or unsuccessful problem detection.

The next section describes the disturbances that trigger problem detection. Following that, we examine the way people make sense of these disturbances—the process of problem detection. Next, we discuss three of the primary factors that affect the detection of subtle problems, in which the cues are muted or ambiguous. These are the most challenging instances for problem detection and often require a reframing of the situation. Finally, we suggest some directions for future research.

5 The disturbances that trigger problem detection

In the incidents we reviewed, people did not receive cues and perform inferences in order to determine if a problem had arisen. Cues are not primitive events—they are constructions generated by people trying to understand situations. People can articulate the significant evidence as cues, but this process of abstracting cues often introduces distortions such as oversimplification (Feltovich et al. 2001). Furthermore, cues are only “objective” in a limited sense. Although they can sometimes be identified with objective data values, they are not purely “input” to be processed as if recognition is solely bottom-up. Rather, the knowledge and expectancies a person has will determine what counts as a cue and whether it will be noticed. (See Mack 2003, for a discussion of inattentive blindness, the phenomenon that people may not notice a stimulus even if they are looking directly at it, if they are attending to something else.)

Table 1 describes the relationship between faults, symptoms, and sensors. Typically, problems can be considered as faults, which are events that threaten to block an intended outcome. However, we do not directly

Table 1 Situational conditions that signal the existence of a problem

Fault →	Symptoms →	Sensors
Routine → deteriorating	Speed of change	Completeness
Recovering → deteriorating	Number and variety	Placement
Single fault	Trajectory	Sensitivity
Multiple faults	Bifurcation points	Update rate
Potential fault	Absence of an event	Cost of use
Reduced margin of safety		Credibility
Weakness in a plan		Turbulence of
Weakness in an action		the background

perceive faults. We notice the disturbances they produce—their symptoms, which we experience as the cues that alert us to the existence of a fault. Whether we notice the symptoms depends on several factors, including the sensors that register the symptoms. The sensor data can be direct (such as visual cues, like the first signs of smoke coming from under the eaves of a building that is starting to catch fire) or indirect (such as a fuel gauge showing a rapid rate of fuel depletion). We have to attend to the sensor data and appreciate how they can be signaling a symptom, in order to suspect that a fault may have occurred. So the faults are signaled as symptoms, and these are noticed if the sensors are appropriately configured.

Faults: A fault can be a shift from a *routine* situation to a *deteriorating* one (e.g., an automobile coming towards you starts to swerve into your lane), versus a shift from a *recovering* situation to a *deteriorating* one (e.g., a patient who seemed to be successfully treated for cancer has a reoccurrence).

The situation can have a *single* fault or it can have *multiple* faults. Multiple faults create a major difficulty in problem diagnosis. They complicate problem detection in that all symptoms may be attributed to the first fault detected, so that another, possibly more damaging fault or fault interactions, may go undetected (e.g., De Keyser and Woods 1993; Woods 1994).

A problem can be detected even in the absence of a fault, if the situation has a *potential* fault. In many cases, the problem is a reduced margin of safety—the person has moved outside the “field of safe travel” (Gibson and Crooks 1938). For example, an increase in wind velocity creates a higher risk for wildland firefighters even though they may not be threatened by any flames. A weakness in a plan is one type of reduced margin of safety. A safeguard that does not materialize is another way that the margin of safety can be reduced, or a discovery that the expected resources will not be available in the quantity and timeframe that was planned.

In the incidents we reviewed, we found many cases where experienced decision makers gauged the riskiness of their own actions. The problems they were sensing were not in the situation (a passive recognition) but within their actions and skills (a projection of actions). The risk potential of actions includes the possibility of

unintended consequences (e.g., firefighters moving to a safe area might find that they were then out of radio contact), and negative consequences (e.g., a surgeon trying to grasp a gallbladder may perforate it instead). A wildland firefighter commented, “If things aren’t going quite as fast as expected, I set an egg timer in my mind and give it a little longer and if the conditions continue to deteriorate, I decide it’s time to withdraw.” One surgeon stated that, “If we haven’t positively identified the cystic duct in another five minutes, we’ll open.” Both comments show sensitivity to the risks in the situations.

Symptoms: Faults, and the disturbances they produce, are experienced as cues, evidence, or symptoms, rather than being directly perceived (as shown in the middle column of Table 1). Manifestation of faults can vary in the *time course of the change*. The change may take place in seconds, hours, or decades. The changes can be sudden, making it easier to detect them. In less than a second, a driving hazard can appear. Alternately, Perrow (1984) provides examples of mining operations and dams where the problem had developed slowly over many years. The time course includes the suddenness of onset of the symptoms. Sometimes, the situation deteriorates in a steady fashion. For instance, in the “going sour” pattern discussed by Cook et al. (1991) and Xiao (1994), the situation is slowly deteriorating but goes unnoticed because each symptom considered in isolation does not signify that a problem exists. At other times there is a rapid deterioration that is more easily spotted. In addition, acceleration cues (changes in the rate of change) are important indicators. We may speculate that Cowan’s (1986) account is designed for zero-order tracking situations, where the change is from an expected to a discrepant state. In contrast, many of the cues in the cases we reviewed involved first and second-order tracking—velocity, acceleration, and even changes in acceleration were important indicators.

The *number and variety* of symptoms can vary, from a single dominant symptom to a set of multiple symptoms. In case 1, presented above, the initial symptom of mottled skin was enough to trigger a heightened alertness, but needed to be connected to several other symptoms in order to result in problem detection.

The *trajectory* can be important. The difference between a safe and an unsafe trajectory is fairly clear at the end but by then there may be insufficient time to make corrections. For example, in case 1, the experienced nurse was trying to infer the trajectory at a very early stage, when the departure from normality was barely noticeable. She was trying to find out how the infant was doing over the last few hours, to see if the symptoms were constant or were getting worse.

A *bifurcation point* can be informative. Within the framework of the chaos theory (e.g., Gleick 1987), a bifurcation point represents an unstable, temporary state that can evolve into one of several stable states. Skilled weather forecasters try to identify these bifurcation points in order to define “the problem of the day” that will need to be closely monitored. Bifurcation

points are most easily identified in hindsight. For weather patterns, the early indicators may primarily be a high variability of pressure and temperature readings. Bifurcation points can be important evidence that a system presumed stable is not. Someone who can detect bifurcation points can be monitoring a situation and be more prepared for events than someone who has to wait until the signs of danger are clear to everyone.

Sometimes, an important data point is the *absence* of an event (Christoffersen et al. 2001). Expertise is needed to notice these “negative” events, which usually are manifested as the violation of an expectancy. In the following example, a critical data point is the lack of radio confirmations of a reported Exocet missile launch.

5.1 Case 3: The inbound Exocet

During Operation Desert Storm, a US Naval Commander on board an AEGIS cruiser received a report that the Iraqis had fired an Exocet missile in his general direction. He quickly prepared his crew to take defensive actions. However, he noted that none of the other ships in the area was generating messages about the Exocet missile. If a missile had been fired, all the ships in its general vicinity should have been detecting it and alerting others. He hypothesized that the report of the missile was probably inaccurate. He maintained a defensive posture, but shifted his primary attention to other matters. His assessment was correct—there was no Exocet missile.

Regardless of the nature of the symptoms, they have to be viewed against a *background*. If the background is very noisy, with lots of potentially relevant signals and distractors, then a person may not notice the symptom because it is not sufficiently discriminable.

Sensors: As shown in Table 1, the disturbances created by the fault are perceived through the use of sensors, both direct (e.g., visual inspection) and indirect (e.g., computer displays of data). Problem detection depends in part on the adequacy of these sensors, and on our understanding of how the sensors work. Mumaw et al. (2000) showed that the operators of a nuclear power plant cannot simply infer the plant status from the displays in the control room. The operators also need a good mental model of the sensors in the plant, and have to be aware of the conditions in the plant itself, including repairs going on and sensors that are malfunctioning.

The sensor system can vary along a number of dimensions:

The coverage can vary in *completeness*. The number of sensors may be insufficient. The *placement* of sensors can be inadequate. (Thus, the absence of a temperature probe at a key point in a petrochemical processing cycle can deprive the operator of an early sign of trouble.)

The *sensitivity* of sensors can vary. A temperature probe may not register temperatures above 150°F and, therefore, cannot be used to discover that the actual

temperature has climbed to 500°F. In addition, valuable time can be lost if a teammate or an automated system detects the early signs of danger but does not publicly announce them. The early signs are thereby masked, so that a slow-to-develop problem is transformed into a no-warning emergency. A high false alarm rate can make it too easy for the person to dismiss symptoms of an actual upset. In our interviews with the laparoscopic surgeons, they commented on the lack of a “feel” for using their instruments, compared to open surgery, showing that they missed this type of sensor.

The *update rate* of the sensors can vary. A slow update or refresh rate can make it difficult or impossible to gauge trajectories early in the cycle. The issue of update rate did not emerge for the surgeons we interviewed. However, for the firefighters we interviewed it was an important feature (e.g., the speed of learning about weather changes, the delays in receiving radio communications).

The sensors may be *too costly* to use. The cost may come in the form of effort and risk (in terms of the safety of the person and the system) for the time needed to operate the sensor and interpret the data. Another type of cost is the ease of adjustment, which affects effort. In our interviews, cost of use was not relevant for the surgeons because the visual display was immediately present, and tactile feedback was eliminated. However, other data collection activities, such as stopping to check for vital signs and performing tests in the middle of the surgery, clearly carried costs in terms of the delay that was required to gather the information. For the firefighters we interviewed, radio discipline is needed to maintain available channels for critical messages.

The *credibility* of the sensor itself can add noise to the system. Credibility is affected by factors such as the perceived reliability of the sensor. Uncertainty about the sensor’s sensitivity also affects credibility, as does inconsistency between sensor data (Schmitt and Klein 1996). A sensor that sometimes malfunctions may not be examined, or its reading can easily be explained away even if it is accurately signaling an anomaly. The *history* of sensor data is important for judging its credibility. The history may be available or it may be inaccessible. In many settings it is important to know how the data were collected, and who were the people doing the collecting. In case 1, the experienced nurse took into account that the infant’s lethargy was apparent even to a relatively inexperienced nurse.

The *turbulence of the background* can change the nature of problem detection. The symptoms of the underlying fault emerge against a backdrop. Operational settings are typically data rich and noisy. Many data elements are present that could be relevant to the problem solver (Woods 1995). There are a large number of data channels (either through sensors, human reports, or direct perception) and the signals on these channels usually are changing. The raw values are rarely constant even when the system is stable and normal. The default case is detecting emerging signs of trouble against a

dynamic background of signals rather than detecting a change from a quiescent, stable, or static background. The noisiness of the background makes it easy to miss symptoms or to explain them away as part of a different pattern (e.g., Roth et al. 1992).

6 Problem detection as a sensemaking activity

The variables shown in Table 1, and examples such as case 1 (*An experienced versus an inexperienced nurse*) and case 3 (*The Inbound Exocet*), suggest a very different account of problem detection than Cowan put forward. Cowan presented a discrepancy accumulation model, in which small cues and signs that were readily perceived added up until they passed a threshold for responding. In contrast, we see the challenge of problem detection as appreciating the significance of data elements in the first place. In order to call a data element a symptom, people already have to appreciate its meaning. For Cowan, problem detection centered around the gap between what was wanted and what was happening; when this gap got large enough a person would perceive that a problem had arisen. In the incidents we studied, problem detection stemmed from the realization that the actual situation was ominously different from the one that the person initially believed.

A considerable amount of expertise is needed in order to make sense of the data received through sensors (also see Klein and Hoffman 1993). Case 1 illustrates the importance of expertise for spotting trouble early on. Note that in incidents such as case 1, the person with experience is not passively receiving sensor data. Instead, the person is trying to make sense of data elements, and is relying on expertise to distinguish actual anomalies from transient fluctuations.

Therefore, we see problem detection as a form of sensemaking. Recently, Klein et al. (2004) have described a data/frame model of sensemaking. The data/frame model asserts that data are used to construct a frame (a story or script or schema) that accounts for the data and guides the search for additional data. At the same time, the frame a person is using to understand events will determine what counts as data. Both activities occur in parallel, the data generating the frame, and the frame defining what counts as data.

The data/frame model also distinguishes between different types of sensemaking activities: elaborating an existing frame, questioning a frame, preserving a frame by explaining away the anomalies, comparing alternate frames to gauge which is more accurate, and reframing to replace an existing frame with a better one. One of the reasons that a function such as sensemaking seems so amorphous and difficult to describe is that it can take all of these different forms. The nature of sensemaking is different depending on whether a person is elaborating or questioning a frame, preserving it or replacing it.

Within the data/frame model of sensemaking, problem detection is about questioning a frame in the first

place—becoming suspicious that the way events are being interpreted is incomplete and perhaps incorrect.

Problem detection also involves other aspects of sensemaking, especially reframing the events, and also preserving a frame by explaining away discrepancies. However, with regard to problem detection, the critical node in the data/frame model is the initial doubt about the way the events are being framed.

6.1 The basis for questioning a frame

Direct contradiction of the frame: The most straightforward basis for questioning a frame is when a salient data element clearly contradicts the person's frame. Case 2 (*Going for the Feint*) illustrated this process. Once the commander realized that a raid of 30 aircraft was approaching he could see that the initial attack was just a diversion, and not the first wave. We can call such a data element a "framebreaker." In many cases, a person simply replaces a mistaken frame with a more accurate one upon being confronted with a framebreaker.

However, there can be complications. De Keyser and Woods (1993) have described the phenomenon of fixation in which a person holds onto a mistaken frame. Feltovich et al. (2001), studying pediatric cardiologists in a garden path scenario, documented the knowledge shields that people use to explain away discrepancies that signal that they are mistaken in the way they are framing the situation. We speculate that the more explaining away people do to preserve their frame, the more resistant they will be to a framebreaker.

Accumulation of discrepancies: A second way in which a person might question a frame is when a set of small discrepancies passes a threshold, as Cowan speculated. According to Cowan, this process should be a fairly linear accumulation of doubts. However, we suggest just the opposite—if the initial small discrepancies have failed to alert a person it is likely that the person is explaining them away, as discussed above. Therefore, the accumulation of discrepancies should have a non-linear effect, with greater commitment to a frame until a point is reached where the frame is suddenly abandoned and replaced by a better one.

Detecting subtle anomalies: A third way in which a person might question a frame is when he or she notices the implications of a subtle cue. The incidents we chose to study primarily fell into this category because these pose the greatest challenge to problem detection. Table 1 lists the reasons why a symptom or cue can be subtle and difficult to notice. The speed of change can be very slow and gradual. The symptom/cue may not be very diagnostic. The difference in trajectory may be very small. The symptom/cue may be the absence of an event. A noisy background may obscure the symptom/cue. The anomaly may be diffuse, in that a range of small, seemingly innocuous clues must be integrated in order to see what is going wrong. A fault in which prior cycles of breakdown and recovery may contribute to the noisiness

that makes it hard to realize that the current cycle is more serious than the preceding ones. A situation with multiple faults will add to the person's confusion. A fault that is a reduced margin of safety may be very hard to gauge, particularly for novices.

Figure 1 adapts the data/frame model of sensemaking, described by Klein et al. (2004), to emphasize the aspects of sensemaking that affect problem detection.

Figure 1 addresses the most challenging aspect of problem detection, realizing that subtle cues are indicating that something may be going wrong. (Other aspects of problem detection include a highly salient cue, as in case 2, or an accumulation of data that pass a threshold.) Table 1 explains how relevant cues can be masked or obscured. As a result, these potential cues are often ignored; other times, they are noticed but explained away so that the initial frame is preserved; and sometimes, they alert decision makers to the possibility of a problem so that they begin to question the frame.

Some of the critical factors that determine whether the subtle cues will be noticed are the degree of expertise of the decision maker, the stance taken, and the influence of stance and expertise to direct attention to the relevant data elements. We discuss these factors in the next section.

7 Factors affecting problem detection

The incidents we studied revealed a number of variables that determined whether someone would notice a problem. Figure 1 shows three of the factors that were most common and with the most impact: expertise, stance, and attention management.

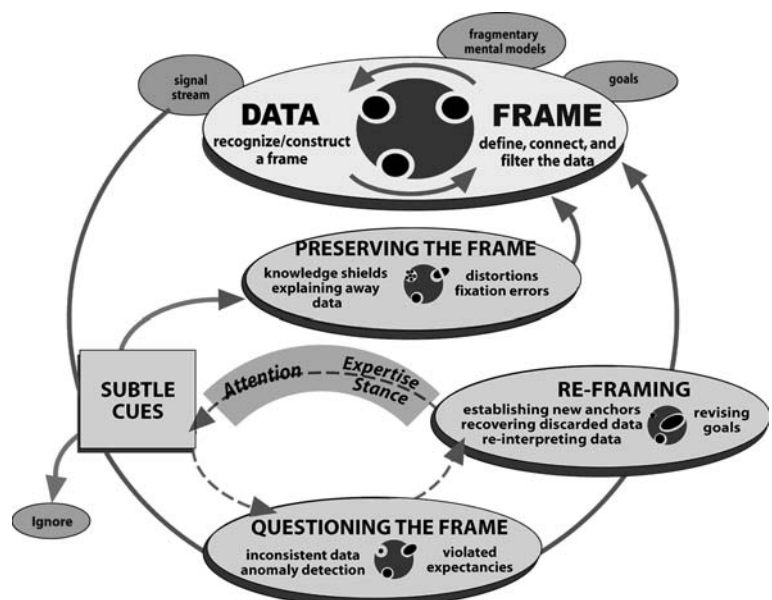
Expertise: Expertise is obviously an advantage in being able to detect problems quickly and accurately, and to question a conceptual frame that may be mis-

taken. Expertise can influence problem detection in many ways. Some of the most important are the perceptual and conceptual ability to notice subtle signs, the ability to use expectancies, the sophistication of mental models, and the experience base that provides a sense of typicality.

The ability to perceive *subtle complexes of signs* and to identify these as cues and patterns is an important aspect of expertise. This includes being able to make fine discriminations. A skilled kayaker can read the water ahead, the hydrotopography that can alert a kayaker to the presence of a submerged rock formation. People need expertise to infer the data elements and evidence from the sensor data that are available. By subtle signs, we mean indicators that most people would not notice. Certainly, we would expect people with more experience to have more accurate signal detection as in the nurse's ability to detect "mottled skin" in case 1 (*An experienced versus an inexperienced nurse*). However, in most of the operational cases, spotting subtle cues was not simply a matter of discriminability. Often, the skill involved detecting a covariation in order to recognize a pattern. For example, the orientee in case 1 could probably have made the distinction between normal and atypical skin color. The failure was not just at the perceptual level. With experience comes a larger set of triggers and alarms; and a larger repertoire of leading indicators that stimulate alertness.

The ability to *generate expectancies* is also important. Skilled personnel can construct mental simulations (Klein and Crandall 1995) that allow them to generate explanations for events that have occurred, tying them together in a story. The skill here is based on a rich causal framework, and richer mental model of situations (e.g., Gentner and Stevens 1983; Johnson-Laird 1983). The expectancies are critical for detecting "negative" cues—events that did not happen, but were supposed to

Fig. 1 Problem detection as a form of sensemaking



occur. Violations of expectancies are important indicators of problems. Moreover, the expectancies create a focus on the aspects to monitor in the future (Christoffersen and Woods 2003). Also, they help a person to better judge urgency—the need to respond quickly rather than waiting to see how things will develop. Also, expectancies may increase suspiciousness, information seeking, and preventative actions. On the other hand, expectations may be erroneous, as when the false alarm rate is high and decision makers disregard legitimate signals.

The person's *mental model* enters into this process in many ways. For example, the mental model of how the sensors work will help a person judge when to trust a sensor and when to seek more information. A novice might begin by believing everything the system said, and then, after running into difficulties, decide never to trust it. A skilled operator is likely to know the conditions under which the system can generate misleading readings, and the causes of those readings. The person's mental model will also affect how he/she interprets the situation. Mumaw et al. (2000) studied operator monitoring of nuclear power plants, and found that the plant status was always changing as equipment failures accumulated. These failures did not compromise the safety of the operations, but many of them could only be repaired during a unit shut down. Therefore, the operators needed to develop a situation model that incorporated the malfunctions, in order to interpret signals.

In our review of critical incidents from previous research, we found a number of examples where the mental model of the sensors and equipment entered into the problem detection process. Many of these examples came from the Neonatal Intensive Care Unit (NICU) research. In the following example, the nurse's mental model made her worry about a baby's blood sugar level. The objective data showed that the blood sugar was fine; however, the nurse explained away these data and held onto her original suspicion, which was confirmed.

7.1 Case 4: The baby with low blood sugar

I was working in transitional when a large for gestational age (LGA) infant was brought in. The infant had been with her mother for almost 2 hours, which was an hour longer than normal procedure. By the time the baby was brought over it was pale and sweating a little bit. When she cried or made voluntary movements her hand and/or foot would shake. The results of the first dextrose stick indicated that her blood sugar was somewhere between 45 and 90. Anything less than 45 is considered to be critical and would call for the lab to do a real blood sugar test. I did not trust the results of the dextrose stick because the baby was shaking, diaphoretic, and pale. She looked like a baby whose blood sugar was low. She was big and big babies tend to have this problem more than regular size babies. I repeated the dextrose stick several times thinking each time that

the results were inaccurate. I worried that either the dextrose sticks were not working or that I had done something wrong. I wondered whether I left the blood on the stick long enough, or maybe I washed it off with hot water instead of cold. Finally, someone from the laboratory was on the unit for something else and I asked them if they would draw a lab test on this baby. I had a feeling that the baby had low blood sugar and if this continued for a prolonged period of time it could lead ultimately to problems such as respiratory distress. (The results of the lab test indicated that the baby's blood sugar was 25, which is very low.)

Olson and Sarter (2001) studied the phenomenon of "management by consent" as employed by airline pilots in simulated scenarios. Automated flight deck systems generated recommendations, and the pilots had to accept or reject these recommendations. The recommendations made by the automated support system led to conflicts with safe operations. However, the pilots were often unable to detect these conflicts because they could not anticipate how the recommendations would affect future configurations.

A sense of *typicality* is important for detecting problems. With experience comes the evolution of prototypes to allow rapid categorization of commonly occurring events. This judgment of typicality is a baseline for detecting anomalies (which are exceptions), violations of the mental models, and patterns and expectancies.

Experience can sometimes *interfere* with problem detection. Thus, De Keyser and Woods (1993) described how a person might fixate on the initial explanation and explain away the anomaly. This type of fixation may be more commonly found among people with expertise, who are better equipped to explain away inconvenient data. De Keyser and Woods characterized this kind of failure of problem detection, based on studies of operator problem solving in both real and simulated emergencies in process control settings such as nuclear power plants and steel mills. When difficulties arose, they tended to be associated with difficulties in revising an assessment as the evidence changed. The people involved tended to hold on to their interpretation of the situation, an interpretation that was correct when first formed or was at least plausible given the limited evidence available early in the episode, despite new evidence that the situation had changed or differed from their assessment. As opportunities to revise occurred, these operators were fixated on their interpretation of the situation, discounting or rationalizing away discrepant evidence. Perrow (1984) has referred to this as a *de minimus* error. Often, problems were detected and diagnosed accurately by new personnel, who entered the situation after the additional data had been obtained, and had not been part of the mindset that was blinding people to the problem (Woods et al. 1987).

Stance: Stance is the orientation the person has to the situation (Chow et al. 2000). The stance can range from denial that anything could go wrong, to a positive

“can-do” attitude that is confident of being able to overcome difficulties, to an alert attitude that expects some serious problems might arise, to a level of hysteria that over-reacts to minor signs and transient signals. Cowan (1986), while not using the term “stance,” suggested that an individual’s “task-role schemas” can direct attention, and that the level of arousal affects an individual’s readiness to respond (i.e., to seek clarification for anomalies). Stance is affected by a person’s level of general alertness, level of suspicion, emotional status, and so forth. The level of *general alertness* is conditioned by factors such as fatigue, and the degree of distraction from competing tasks and workload.

Sometimes people adopt a highly *suspicious* stance or attitude, as when a skilled weather forecaster comes in to work searching for the problem of the day, which is the unsettled part of the scene that will need to be closely monitored. Less-skilled forecasters often take the stance of trying to make the best estimate without this type of active searching (Pliske et al. 2004). This active searching seems linked to an engaged attitude.

In our review of the problem detection incidents with nurses in the NICU, we found that in half the cases the nurses were in a vigilant, alert state that helped them see the significance of early cues. Thus, in case 1, the experienced nurse was looking for indications of babies who were falling ill. She had seen enough babies ‘go sour’ that she was actively searching for trouble spots, in comparison to the orientee who was passively doing her job of monitoring and recording. The nurse in case 4 (*The baby with low blood sugar*) was uncomfortable with the results of the test because the baby’s condition marked it as being at risk for low blood sugar, and because of the way the baby looked. The way this nurse carried out her work was to be on the lookout for these anomalies.

Emotional status would include the current level of anxiety and the attention paid to anxiety cues, along with other cues from other emotions. Anxiety is an important cue for detecting problems, as shown by Bechara et al. (1997). Normal participants showed psychophysiological anxiety reactions to risky situations before they were cognitively aware of the risk, whereas brain-damaged participants, lacking these psychophysiological mechanisms, were much less able to detect and react to risk. The degree to which a person is sensitive to these cues could impact the speed and accuracy of problem detection. Similarly, a person who is already anxious, and is monitoring general signs of anxiety, fear and other emotional states might have difficulty with problem detection.

Attention management: One of the ways that expertise and stance are manifested is through attention—what is ignored, what is monitored, what is scanned for. Research on inattentive blindness (Mack 2003) shows that people may not even notice stimuli right in their field of view, if they are not expecting to see them or if they are irrelevant to the task at hand.

The function of managing attention includes handling the configuration of sensors, as shown in Table 1.

The credibility of sensors e.g., the nurse’s lack of trust in the standard test in case 4 (*The baby with low blood sugar*), the completeness and geometry of coverage, the sensitivity and update rate, are all part of the management task, balanced against the cost of using different sensors.

Figure 1 shows that detecting the significance of subtle cues that signal a problem is also affected by the way a person reframes a situation. The process of reframing is central to sensemaking, and to problem detection.

8 Problem detection as reframing the situation

Cowan (1986) asserted that problem detection occurred when discrepancies accumulated between what was being observed, and what was desired, until the discrepancies pass some threshold and become noticed. This type of account does not address the difficulty of noticing a discrepancy in the first place, which seems to be the heart of problem detection in many situations.

We disagree with Cowan’s formulation. Our view is that problem detection is about discrepancies between what is observed and what is expected, more than discrepancies between observed and desired states (Woods et al. 2002). We agree with Cowan that a person has a problem if events are taking the wrong turn. However, one of the things that make problem detection so difficult is the need to generate expectancies and to notice where these are violated, and to estimate a trend for the violation to continue and increase. The kinds of expertise needed here are different from the kinds needed in Cowan’s (1986) framework.

In our view, the cues or anomalies that trigger problem detection are not automatically given by the situation. They are constructed, inferred, and hypothesized. Moreover, problem detection is not a matter of exceeding a threshold for discrepancies. It requires people to reframe the way they understand the situation. We are not seeing problem detection as a trigger for sensemaking, but as an aspect of sensemaking, and even as a microcosm of sensemaking (Klein et al. 2004; Weick 1995).

In some incidents, such as case 2, a disturbance was so salient that it seemed to drive the problem detection process. In other incidents, such as case 1, many of the signs of disturbance were only clear to the person who was already worried about the situation. Some of the anomalies/discrepancies will only become clear once the problem is noticed, and yet these indicators are the basis for noticing that there is a problem.

This circularity raises the question of what comes first, the indications that trigger problem detection, or the detection of the problem, which conditions the interpretation of the indicators? This dilemma is precisely the problem of meaning recognition, otherwise known as the Höfding Problem (Höfding 1889). How can you recognize something before you know what it is

that you are recognizing? The act of recognition presupposes an act of recognition—this conundrum comes from thinking of it as a purely bottom-up process. Sometimes, the cues may come first, as in the Cowan model, but there also are cases in which the indicators and the reinterpretation appear to occur together. In Fig. 1, questioning the frame leads a person to reframe the situation, but a person needs to be reframing the situation in order to appreciate the significance of the subtle cues. These activities, questioning and reframing, are shown as separate in Fig. 1, but they may actually be the same activity.

Therefore, we view problem detection as a process of reframing or reconceptualization. The process of reframing is specifically required for noticing subtle cues, rather than highly salient cues such as case 2 (*Going for the faint*). Case 1 (*An experienced versus an inexperienced nurse*) is a vivid example of how reframing permits a decision maker to see a pattern, to notice connections, to look for diagnostic data elements.

We are suggesting that in some subtle cases, there may not be a meaningful or even legitimate distinction between cue recognition and problem detection. We can use different terms for them, and we can prepare diagrams in which they appear in different boxes (along with arrows to show iterations). If we try to build computer simulations, we could invoke different sub-routines. But at a psychological level, the distinction may be misleading. The early detection of symptoms and fault indicators may be the same thing as reconceiving a situation as one that is problematic.

Examples such as case 1 show how problem detection can involve a shift in the practitioner's conception of the situation. The experienced nurse did not look at the baby's skin, take in the observation that the color was off and the skin was mottled, think about what that might signify, and decide that she should collect some more information. She saw that the skin color was off and was mottled, and that perception was the same as realizing that the baby might be in trouble. In contrast, the less-experienced nurse, who was not able to make judgments about skin changes and had to rely on unambiguous cues such as temperature, missed the trends and subtle cues because her conception of the baby's condition had not changed.

It could be argued that case 1 shows how a large number of cues resulted in problem detection, following Cowan's model. Certainly, there was evidence for the onset of sepsis (although not so much as to alert the trainee). However, the instructor was detecting a problem from the outset, upon noticing the baby's skin color. The color and mottled nature of the skin was seen as an anomaly, to a nurse whose stance was to look for anomalies, and resulted in more directed information seeking. It has become a way that experienced nurses look at infants in the NICU. Also, because this is how the experienced nurse in case 1 has learned to look at babies, she was prepared to spot the initial signs. She already had available an alternate conception—a baby

with sepsis—and so she could reconceptualize the situation in detecting the initial cues.

From this perspective, failures of problem detection are not so much failures to detect an indicator, but rather they are failures to reconceive or redefine the situation. The trainee in case 1 could likely have made a distinction between different skin colors and different degrees of mottling for the babies. The cues were not below sensory thresholds. Rather, the trainee did not have an alternative conceptualization available to guide her perceptions.

In failures of problem detection, monitoring and stance stay on the current line of activity. In the cases we reviewed, problem detection as nascent reconceptualization was revealed in the way practitioners reassessed whether they should shift their account of events because they may be facing a new type of situation. Hence, the positive role of anxiety in problem detection can be seen as helping to question whether the current assessment and the current line of activity are still appropriate (e.g., case 4 *The baby with low blood sugar*).

In describing sensemaking, Weick (1995) notes that it is an activity in which many possible meanings may need to be synthesized, because many different projects are under way at the time reflection takes place "...the problem is that there are too many meanings, not too few... the problem is confusion, not ignorance." (p. 27) In our view of problem detection as reconception, the sensemaking that Weick has discussed is not reserved as a stage within problem detection. Rather, it spans the problem detection process. The detection of a critical indicator or pattern is a revision in the understanding of the situation. More deliberate sensemaking can follow, but the person has already reinterpreted the situation.

8.1 Various forms of problem detection

Beyond the description of problem detection as reconceptualization, we identified a range of different strategies. The variability suggests that it will be unproductive to attempt to specify a general problem detection strategy. Problem detection is different when responding to a salient message than it is in noticing a pattern, or in bowing to the weight of accumulated evidence, or in breaking free from a fixation that has been deflecting contrary evidence.

In some incidents, a significant amount of expertise was needed to notice the critical data elements. The skill seemed to require a great deal of perceptual learning. For some subtle cases, reconceptualization is required to spot the anomaly in the first place. In some cases, we speculate that there is something like an antibody reaction. The weak cues elicit a strong reaction, presumably due to prior experience.

In some cases, the indication of a disturbance is very clear, so the data would drive the detection. Many of these incidents were trivial, such as case 2 (*Going for the faint*), but some involved the de minimus tactic of

preserving an incorrect picture of the situation by explaining away anomalies. In these cases, people only reconceptualized the situation when confronted with incontrovertible evidence.

In some incidents, the key was noticing inconsistencies between data elements. This strategy was conceptual rather than perceptual. Case 4 (*The baby with low blood sugar*) shows how a nurse was troubled by inconsistencies between the objective data from the lab work, and her own impressions of the infant's blood sugar level. Case 5 is another example of how problem detection can depend on noticing inconsistencies.

8.2 Case 5: Tremors or seizures?

I was working with a very little infant (25- to 26-week-old) that nobody really expected to live. He was my primary so I worked with him every night. When he was about 2–3 days old I noticed real fine tremors in his extremities on one side. He had these tremors fairly regularly throughout the shift and I charted them as tremor activity. I went to the charge nurse and told her that I thought the baby was having seizures. The doctors did not think they were seizures. The next night he was still having fine tremors and it seemed that they were increasing in frequency and strength. Each time he tremored his oxygen saturation dropped. He also was doing some tongue thrusting but it was so mild it looked like a suck reflex; but 25-week-old babies do not normally have a suck reflex.

By the third night I convinced one of the physicians that he might actually be having seizures. Diagnostic tests were ordered and as it turned out he had a grade 4 bleed. He was put on phenobarbital, the seizures were resolved, and he eventually went home with no damage.

Sometimes the trigger for problem detection was that several different anomalies were seen. One could argue that perhaps problem detection in these cases depended on the accumulation of evidence but it is also possible that the different indicators served as converging evidence and the person was less worried about being fooled by a transient data element that might prove to be unreliable. In addition, the mere number of anomalies seemed less important than the fact that they fit a pattern.

This last strategy that we have described is consistent with the anomaly accumulation model posited by Cowan (1986). However, Cowan's account does not address features of the other strategies: the difficulty of noticing a discrepancy, seeing patterns of cues, or any of the sensemaking needed to appreciate the significance of an anomaly. None of the cases we studied fit a framework of accumulating discrepancies. In virtually every case, expertise was needed in order to define, detect and interpret cues. In fact, the concept of a cue often becomes murky in natural situations. In case 1, the instructor noted how the baby's temperature had been falling over time. Was the cue the discrete temperature

readings, interpreted to show a pattern of change? Or was the cue the change itself? The concept of "cue" is a construction across multiple relationships, not a primitive feature of situations.

9 Directions for future research

One reason to try to describe a complex phenomenon is to raise questions for further investigations. We offer the following suggestions.

More intensive empirical studies, including opportunities to observe situations in which problem detection occurs. The interviews we conducted in this project cover only two domains, and fewer than ten incidents, to go along with a post hoc review of data collected in other studies. Future investigations should cover a broader range of domains. In order to facilitate such research into problem detection, we will need a standardized methodology for defining problem detection incidents, for collecting the relevant contextual data through observations, interviews, and post-incident reviews (similar to the framework for documenting possible cases of fixation in De Keyser and Woods 1993).

Effects of stance, expertise and attention management. We have speculated that these variables affect the success of problem detection. Therefore, by manipulating each of them researchers should be able to increase and decrease problem detection performance.

Domain-specific failures. It is possible that a specific domain has characteristic barriers to problem detection, such as types of stance or masking, common limitations of the sensors, or typical types of disturbances. For example, the problem detection failures in NASA's Challenger accident share many features with the Columbia accident. Perhaps, we can characterize domains in order to identify frequent types of faults, frequent shortcomings of sensors, typical types of problem detection strategies, and frequent types of problem detection breakdowns.

Nonlinear resistance. Research on fixation and garden path effects and knowledge shields suggest that under some conditions, the more evidence people explain away the more difficult it is to see the significance of new data elements that contradict a dominant hypothesis, until a point is reached where the person breaks free of the fixation. Thus, the impact of disconfirming evidence is not a smooth increase in the probability of rejecting the hypothesis, but rather a disregard for the evidence and sometimes even a strengthening of the original hypothesis, until the breaking point is reached. Can we reliably demonstrate this effect and determine the conditions under which it occurs? The prediction is that the effect of a "framebreaker," a conclusive piece of evidence, will be diminished if it is preceded by a series of minor anomalies that can be explained away.

Human-automation teamwork. Layton et al. (1994) and Guerlain et al. (1999) have demonstrated the conditions leading to both poor and to effective couplings

of machine advice. When the machine generates potential solutions for people to review or critique, Layton et al. found that people missed flaws in the machine's recommendation. Guerlain et al. tested alternative human-automation strategies where the machine's analysis is embedded in a human-computer support system as reminders and critiques of person's process. This architecture enhanced the ability of people to handle difficult cases—cases difficult for the machine alone or the people alone. We hypothesize that if people are handed computer-based appraisals of situations, these people will be slower to detect the early signs of a problem than if they had to build their own picture of what is going on.

Coping with massive amounts of data. Technology allows access to massive amounts of data, which creates data overload problems (Woods et al. 2002). Research is needed to determine if new forms of organizing data into pattern-based visualizations will help or hinder problem detection. If element-based organizations hinder problem detection, pattern-based organizations may be preferred.

Alarms and reminders are being used to direct operators' attention to signals. However, the false alert problem remains a considerable barrier (Sorkin 1988). Even apparently sensitive alerting systems provide little information if false alarms are high (high alerting rate creates low positive predictive value). Displaying analog alerts based on change in likelihood is a promising technique that could aid problem detection (Sorkin et al. 1998).

Developing assessment methods in a domain to identify common barriers and errors regarding problem detection. In our limited investigation, we have seen that the nature of the domain affects the type of barriers to problem detection. Some domains, such as firefighting, pose high degrees of uncertainty about how the situation will develop, whereas other domains such as surgery, pose uncertainty about the consequences of actions that can lead to unintended damage. Before we can identify tactics for training or design to improve problem detection in a domain, we will need to clarify the types of problems that are difficult to detect, the reasons why they are difficult, the common errors, and so forth.

Developing training programs to improve problem detection skills. Several generic methods already exist for problem detection training. Cohen et al. (1998) have developed a "crystal ball" technique to reduce fixation. Klein (1997) describes a "PreMortem" method to identify weaknesses in a plan. Both these methods may be useful to improve problem detection.

Studying team and organizational barriers to problem detection. In most settings, the breakdown of teamwork and the accumulation of organizational inefficiencies may be a greater threat to problem detection than a lack of individual expertise. Engdahl and Keating (1995) examined some problem detection processes at the group and organizational level. It may be useful to expand this line of research.

10 Conclusions

The topic of problem detection has been relatively neglected in the cognitive science literature, despite its importance. A previous account of problem detection (Cowan 1986) described how discrepancies mount until a threshold for detection is triggered. While we agree with many of the factors that Cowan has identified, our review of problem detection cases revealed some major limitations in his description. In natural settings, it is not trivial to notice discrepancies. Often, a person can detect a discrepancy only if that person is prepared to reconceptualize the situation. The critical symptoms may be invisible to someone who is not, at some level, already looking for them. Therefore, the reconception of the situation and the detection of anomalies may be the same psychological activity.

We also found that it is not obvious what should count as a "cue." Many of the cues are so subtle that only an expert would see them, and require an active stance of searching for difficulties. Mental models can also be essential in recognizing cues. Further, what counts as a critical cue depends on the nature of the fault, the nature of the symptoms, the characteristics of the sensors, and the noisiness of the background. This account is different from the stimulus-response or antecedent-production rule type of description in which an unambiguous stimulus triggers a learned reaction. In the tasks driving the incidents we studied, there were no unambiguous "stimuli." Echoing the critique that Dewey (1896) made of the reflex arc concept in psychology, we found that the way the situation was understood, and the repertoire of potential reactions, conditioned the recognition of anomalies.

We have tried to make explicit some of the ways that expertise can be applied. Expertise is in the form of accurate mental models (to generate expectancies and to notice that expectancies have been violated) and skills to make perceptual discriminations and recognize patterns. These abilities enable people to anticipate when the margin for error has become too small, and to detect subtle cues that are often early signs of a problem. We have also described some of the reasons why the mindset of an experienced person can interfere with problem detection.

An additional psychological factor that appears to affect problem detection is stance. This includes workload, fatigue, and task factors that encourage active problem searching versus attempts to explain away anomalies. Both expertise and stance affect the way a person manages attention, which also influences the effectiveness of problem detection.

Our examination of incidents from natural environments suggests that problem detection is not at all straightforward, and that it involves the full range of complexity, such as sorting through noisy backgrounds, and mental simulation, found in most cognitive phenomena. Problem detection is a form of sensemaking,

and appears to require a reconceptualization of a situation.

Klein et al. (2003) have distinguished microcognition (the study of cognitive phenomenon under controlled conditions) with macrocognition—the study of cognitive functions that are required in field settings. Many of the macrocognitive functions are rarely studied in the laboratory. Klein et al. listed problem detection as a leading example of a macrocognitive function that is critical in natural settings despite the lack of experimental scrutiny. In this article, we have tried to explain why problem detection poses such a macrocognitive challenge.

Acknowledgements We would like to thank Marvin Thordsen, Rob Hutton, and Patty McDermott for conducting the interviews with wildland firefighters and laparoscopic surgeons, and reviewing the transcripts of taped interviews with surgeons. We would also like to thank Robert Hoffman and several anonymous reviewers for making many helpful suggestions. Erik Hollnagel challenged our formulation and helped us clarify our ideas. We also appreciate the support and encouragement we have received from Fumiya Tanabe. This project was funded by several agreements with Japan Atomic Energy Research Institute.

References

- Anderson JR (1993) Problem solving and learning. *Am Psychol* 48(1):35–44
- Bechara A, Damasio H, Tranel D, Damasio AR (1997) Deciding advantageously before knowing the advantageous strategy. *Science* 275:1293–1295
- Billings R, Milburn T, Schaalman M (1980) A model of crisis perception: a theoretical and empirical analysis. *Admin Sci Q* 25:300–316
- Chow R, Christoffersen K, Woods DD (2000) A model of communication in support of distributed anomaly response and replanning. In: Proceedings of the XIVth Triennial Congress of the International Ergonomics Association and 44th annual meeting of the human factors and ergonomics society, vol 1, pp 34–37
- Christoffersen K, Woods DD (2003) Making sense of change: extracting events from dynamic process data (Report ERGO-CSEL 01-TR-02). Institute for Ergonomics/Cognitive Systems Engineering Laboratory, Columbus
- Christoffersen K, Woods DD, Blike GT (2001) Extracting event patterns from telemetry data. In: Proceedings of the human factors and ergonomics society 45th annual meeting, Minneapolis/St. Paul, 8–12 October 2001. HFES, Santa Monica, pp 409–413
- Cohen MS, Freeman JT, Thompson B (1998) Critical thinking skills in tactical decision making: a model and a training method. In: Cannon-Bowers J, Salas E (eds) *Making decisions under stress: implications for individual and team training*. APA Press, Washington, pp 155–189
- Cook RI, Woods DD, McDonald JS (1991) Human performance in anesthesia: a corpus of cases. *Cognitive Systems Engineering Laboratory Report*, prepared for the Anesthesia Patient Safety Foundation, April 1991
- Cowan DA (1986) Developing a process model of problem recognition. *Acad Manage Rev* 11(4):763–776
- Crandall B, Calderwood R (1989) Clinical assessment skills of experienced neonatal intensive care nurses (No. Contract 1 R43 NR0191101 for The National Center for Nursing, NIH). Klein Associates, Fairborn
- Crandall B, Getchell-Reiter K (1993) Critical decision method: a technique for eliciting concrete assessment indicators from the “intuition” of NICU nurses. *Adv Nurs Sci* 16(1):42–51
- Davis GA (1973) *Psychology of problem solving*. Basic Books, New York
- De Keyser V, Woods DD (1993) Fixation errors: failures to revise situation assessment in dynamic and risky systems. In: Colombo AG, Saiz de Bustamante A (eds) *Advanced systems in reliability modeling*. Kluwer, Norwell
- Dewey J (1896) The reflex arc concept in psychology. *Psychol Rev* 3:357–370
- Dominguez C (1998) Expertise in laparoscopic surgery: anticipation and affordances. Paper presented at the 4th naturalistic decision making conference, Warrenton, VA
- Dominguez CO, Flach JM, McDermott PL, McKellar DM, Dunn M (2004) The conversion decision in laparoscopic surgery: knowing your limits and limiting your risks. In: Smith K, Shanteau J, Johnson P (eds) *Psychological explorations of competence in decision making*. Cambridge University Press, NY, pp 7–39
- Downs A (1967) *Inside bureaucracy*. Little and Brown, Boston
- Duncker K (1945) On problem solving. In: JF Dashiell (ed) *Psychological monographs*. Translated by LS Lees 58(5, Whole No. 270), pp 1–113
- Engdahl RA, Keating RJ (1995) Strategic problem-solving process: a comprehensive model and the need for more focused attention on consensus seeking. *Organ Develop J* 13(3):80–94
- Feltovich PJ, Coulson RL, Spiro RJ (2001) Learners’ (mis)understanding of important and difficult concepts: a challenge to smart machines in education. In: Forbus KD, Feltovich PJ (eds) *Smart machines in education*. AAAI/MIT Press, Menlo Park
- Flanagan JC (1954) The critical incident technique. *Psychol Bull* 51:327–358
- Forbus KD, de Kleer J (1993) *Building problem solvers*. The MIT Press, Cambridge
- Gentner D, Stevens AL (eds) (1983) *Mental models*. Lawrence Erlbaum Associates, Mahwah
- Gibson JJ, Crooks LE (1938) A theoretical field-analysis of automobile driving. *Am J Psychol* 51:453–471
- Gleick J (1987) *Chaos: making a new science*. Viking Penguin, New York
- Greeno JG, Simon HA (1988) Problem solving and reasoning. In: Atkinson RC, Herrnstein RJ, Lindzey G, Luce RD (eds) *Stevens’ handbook of experimental psychology: vol 2, Learning and cognition*. Wiley, New York
- Guerlain S, Smith PJ, Obradovich J, Rudmann S, Strohm P, Smith JW, Svrbely J, Sachs L (1999) Interactive critiquing as a form of decision support: an empirical evaluation. *Hum Factors* 41(1):72–89
- Hayes JR (1981) *The complete problem solver*. The Franklin Institute Press, Philadelphia
- Höfding H (1889) Ueber Wiederkennen, Association und psychische Activat. *Vierteljahrsschrift für wissenschaftliche Philosophie* 13:420–458
- Hoffman RR, Crandall BW, Shadbolt NR (1998) Use of the critical decision method to elicit expert knowledge: a case study in cognitive task analysis methodology. *Hum Factors* 40(2):254–276
- Johnson-Laird PN (1983) *Mental models: towards a cognitive science of language, inference, and consciousness (Cognitive Science, No. 6)*. Harvard University Press, Cambridge
- Kaempf GL, Klein G, Thordsen ML, Wolf S (1996) Decision making in complex command-and-control environments. *Hum Factors* 38:220–231
- Klein G (1997) Developing expertise in decision making. *Thinking Reasoning* 3(4):337–352
- Klein GA, Crandall BW (1995) The role of mental simulation in naturalistic decision making. In: Hancock P, Flach J, Caird J, Vicente K (eds) *Local applications of the ecological approach to human-machine systems, vol 2*. Lawrence Erlbaum Associates, Mahwah, pp 324–358
- Klein GA, Hoffman R (1993) Seeing the invisible: perceptual/cognitive aspects of expertise. In: Rabinowitz M (ed) *Cognitive science foundations of instruction*. Lawrence Erlbaum Associates, Mahwah, pp 203–226

- Klein G, Hutton R (1995) The innovators: high-impact researchers at the Armstrong Laboratory Human Engineering Division, Armstrong Laboratory (U) (AL/CF-FR-1995-0027). United States Air Force Armstrong Laboratory, Wright Patterson AFB, OH
- Klein GA, Calderwood R, Clinton-Cirocco A (1988) Rapid decision making on the fireground (Technical Report No. DTIC No. AD-A199492, <http://www.dtic.mil>). US Army Research Institute, Alexander
- Klein GA, Calderwood R, MacGregor D (1989) Critical decision method for eliciting knowledge. *IEEE Trans Syst Man Cybern* 19(3):462–472
- Klein G, Ross KG, Moon BM, Klein DE, Hoffman RR, Hollnagel E (2003) Macrocognition. *IEEE Intell Syst* 18(3):81–85
- Klein G, Phillips JK, Rall E, Peluso DA (2004) A data/frame theory of sensemaking (submitted for publication)
- Klinger DW, Militello LG (2002) Designing for performance: a cognitive systems engineering and cognitive task analysis approach to the modification of the AWACS weapons director interface (submitted 10/2001)
- Layton C, Smith PJ, McCoy CE (1994) Design of a cooperative problem-solving system for en-route flight planning: an empirical evaluation. *Hum Factors* 36(1):94–119
- MacCrimmon KR (1973) Managerial decision-making. In: McGuire JW (ed) *Contemporary management: issues and viewpoints*. Prentice-Hall, Englewood Cliffs, pp 445–495
- Mack A (2003) Inattention blindness: looking without seeing. *Curr Dir Psychol Sci* 12(5):180–184
- Mumaw RJ, Roth EM, Vicente KJ, Burns CM (2000) There is more to monitoring a nuclear power plant than meets the eye. *Hum Factors* 42(1):36–55
- Newell A, Simon HA (1972) *Human problem solving*. Prentice-Hall, New Jersey
- Olson WA, Sarter NB (2001) Management by consent in human-machine systems: when and why it breaks down. *Hum Factors* 43(2):255–266
- Perrow C (1984) *Normal accidents: living with high-risk technologies*. Basic Books, New York
- Pliske RM, Crandall B, Klein G (2004) Competence in weather forecasting. In: Smith K, Shanteau J, Johnson P (eds) *Psychological investigations of competence in decision making*. Cambridge University Press, Cambridge, pp 40–70
- Polya G (1957) *How to solve it*, 2nd edn. A Doubleday Anchor Book, Garden City
- Roth EM, Woods DD, Pople HE (1992) Cognitive simulation as a tool for cognitive task analysis. *Ergonomics* 35:1163–1198
- Rubinstein MF (1975) *Patterns of problem solving*. Prentice-Hall, NJ
- Schmitt JF, Klein G (1996) Fighting in the fog: dealing with battlefield uncertainty. *Mar Corps Gaz* 80:62–69
- Schrenk LP (1969) Aiding the decision making—a decision process model. *Ergonomics* 12:543–557
- Smith GF (1989) Defining managerial problems: a framework for prescriptive theorizing. *Manage Sci* 35(8):963–981
- Sorkin RD (1988) Why are people turning off our alarms? *J Acoust Soc Am* 84:1107–1108
- Sorkin RD, Kantowitz BH, Kantowitz SC (1998) Likelihood alarm displays. *Hum Factors* 30:445–459
- Watts-Perotti J, Woods DD (1997) A cognitive analysis of anomaly response in space shuttle mission control (Report CSEL 97-TR-02C). The Ohio State University, Columbus
- Weick K (1995) *Sensemaking in organizations*. Sage Publications, Thousand Oaks
- Wertheimer M (ed) (1959) *Productive thinking*. Harper and Row
- Woods DD (1994) Cognitive demands and activities in dynamic fault management: abduction and disturbance management. In: Stanton N (ed) *Human factors of alarm design*. Taylor and Francis, London
- Woods DD (1995) The alarm problem and directed attention in dynamic fault management. *Ergonomics* 38(11):2371–2393
- Woods DD, O'Brien J, Hanes LF (1987) Human factors challenges in process control: the case of nuclear power plants. In: Salvendy G (ed) *Handbook of human factors/ergonomics*. Wiley, New York
- Woods DD, Patterson ES, Roth EM (2002) Can we ever escape from data overload? A cognitive systems diagnosis. *Cogn Technol Work* 4(1):22–36
- Xiao Y (1994) *Interacting with complex work environment: a field study and a planning model*. PhD thesis, University of Toronto