One of the most powerful knowledge elicitation methods available to Cognitive Task Analysis (CTA) practitioners is to probe actual incidents. People tell us about all kinds of details, challenges, subtle cues, background influences, and strategies that might never come to light in a general interview or a controlled simulation. Skilled decision makers have had many different experiences; that's how they formed their knowledge and honed their skills. Their stories can be a doorway into that experience.

Take weather forecasting as an example. Experienced forecasters have always traded insights using case studies of particular weather events. Case studies appear regularly in journals such as *The Monthly Weather Review* and *Weather & Forecasting*. Senior forecasters love to tell stories about their first successful tornado forecast or the first time a fog forecast "busted." Some of the stories[1] they remember are rich in detail:

It was midwatch. Before the Mobile [Alabama] radar painted it, we knew only that there was a southwest air flow and clouds down over the Gulf south of New Orleans. It was a bad system. It was southwest, about 100 miles south of the mouth of the Mississippi River. It was a big storm cell. It was moving north-northeast. I knew it would hit close by and would affect our area. After one hour I knew it would qualify as a supercell. When it crossed the mouth of the Mississippi, the Weather Channel said, "Look at this supercell!" We'd been looking at it for over an hour. Slidell and Mobile radars were getting good reads on it. I kept extrapolating the track via the NEXRAD radar. I watched it loop by loop. Bad cells tend to turn to the right but they can sometimes turn to the left. If it is upstream of you, you are not going to take your eyes off of it. I knew it would hit at about 3:00 AM. When it got 40–50 miles south-southwest, I realized it would track 20 miles east, right about at Smith Field. I called Smith Field at about 1:00 or 2:00 AM. They had a young forecaster there, just out of [his first duty assignment], and hadn't worked any severe weather. I asked if he was aware of the supercell heading toward him. He said, "....What?!"

Many forecasters can even pull out their "special" files of the records they have kept of interesting and tough cases they experienced. They thrive on the details and have clear awareness of each of the lessons they learned. Once a forecaster gets going on a favorite story, he or she can take an hour or more simply to lay out all the details.

With luck, apprentice forecasters may get a chance to learn from the experiences of the senior practitioners. Here we see, of course, one of the applications for CTA—the lessons learned that are contained in stories also knowledge sharing and training. The lessons learned that are contained in stories also suggest leverage points, perhaps for new decision aids. Managers can also use stories to appreciate what makes their staff members expert and to take that into account in running an organization.

What sorts of things can the CTA researcher find in stories?

- The cues and patterns that experts perceive:

I could see the air pressure falling and knew I could put out a warning for strong winds. [The lightning network] showed a ring of lightning around the "Low" pressure center. This was unusually symmetrical, but showed that the Low was well organized. From a hand plot of buoy data about air pressure, I could plot the front, the Low's position, movement, rate of movement. I did about one plot per hour, about six or eight in all. Enough to know that the warning had to go out and then two or three more plots to show that it really was out there.

- The rules of thumb they have devised:

It was a warm air mass over a cold air mass condition, which trapped the fog. Gulf of Mexico moisture was coming up due to high pressure over the Gulf. The airport was just high enough in elevation to condense the moisture and form fog. The forecasting problem was if and when the fog ceiling would raise enough for flights. [Trainee pilots] needed to fly. And it was midweek so they were busy. We'd look over the [airfield] toward the downtown hotels and use the hotels as ceiling indicators. The downtown is 15 miles away. If you could not see the top of a certain building, you knew the ceiling was 800 feet. We knew from the visibility of the hotel floors what the ceiling was, and when it got up to 800 feet. There were other rules of thumb. If you could not see the airport tower you knew the visibility was less than 3/4 of a mile. You use what you can. The pilots kept bugging me so I had to keep monitoring the situation—satellite loop, visibility, every 5 to 10 minutes, observations out of the area airfields. By 1:00 PM I knew no one would fly.

- The kinds of decisions they have to make:

I came on midwatch duty Saturday evening. The National Hurricane Center (NHC) had Hurricane Georges tracking west-northwest. The computer forecasting models had it going every which way after landfall. The NHC had the wrong track. They were wrong on where the eye of the hurricane was. We could see it on radar. You could see the eye wobble on the satellite image loop and the radar loop. The eye was running in and out and sometimes was defined and sometimes was not. We looked at buoy data every few hours and did our own charts. [The participant in this interview had kept the originals and pulled them out of a file drawer.] The NHC shifted the hurricane track a little to the east out to Gulfport Mississippi, but we were leery about that track. They were still off. The NHC had it shifting northwest to Louisiana, more of a westward track. But we could see it heading due north toward Biloxi. We had to go with the official forecast.

- The features that make decisions tough:

The analysis of the upper atmosphere showed an area of turning winds. My goal was to try to figure out what would happen. If it kept moving through, nothing would happen, but if it didn't, you get caught with your pants down. The wind shift implied that something was happening. A novice would have missed it. The region of maximum wind curvature was at the top of the high pressure ridge so it would not show up so much. The turning of the winds was enough to me. Then you look for support and weigh all the factors. [The cirrus] was not enhancing, but it also was NOT going away. There was moisture at the upper levels. Water vapor imagery showed no slot of dry air associated with the front, implying there was no instability. But I was seeing that there was potential here. There must have been something balancing the wind curvature. The upper-level wind should have changed as it was going down the ridge.

- The features that make cases typical:

It was March. Before arrival [at the weather station] I was skywatching. It was not a blue or gold sunset. I saw cirrus to the southwest, anvil cirrus blowing off the tops. You can see this even though the main clouds might be 100–200 miles away. This confirmed that there was energy out there. There were not enough data yet. We had to query the buoys. This was a textbook case. A stalled front off the Texas coast. You look out to the southwest and if you see any approaching trough, vorticity, or a vorticity maximum, any Low or wave on the front will develop one or two storm systems. It is taught in the School and is discussed in the Local Handbook. But you still need to experience it firsthand a few times. If you get burned once, then you learn.

- The features of rare cases:

Maintained gale-force winds require major storm systems. This is rare. The storm of March 1993 hit western Florida with 112 mph winds. That situation was similar to this one—everything lines up perfectly. But major storms out of this scenario are rare. These were minor storms. I was asking myself, were the Lows intensifying and moving eastward? Intensification would imply a need to upgrade the warning. Would people need to do preparations at the [airfield]? It was not a routine situation since it is not usual to get a heavyweight supercell at midshift. This was a standard scenario in terms of the storm development and dynamics, but not standard in terms of the time of year and time of day of the storm. Fast-moving cold fronts coming from the west usually determine our winter weather—storms and small lines of storms. A big cell developing in a southwest flow is rare for winter in the Gulf region.

These kinds of information are contained in stories that can be elicited in any domain. Military leaders, project managers, nurses, sales personnel, firefighters, even consumers can describe incidents for the CTA researcher to study. We have developed the Critical Decision Method (CDM) to learn from specific incidents (Hoffman, Crandall, and Shadbolt 1998; Klein, Calderwood, and MacGregor 1989). Many CTA researchers use the CDM for conducting incident-based interviews (Blandford and Wong 2004; Ebright et al. 2003; Klein and Armstrong 2004; Militello and Crandall 1999; Omodei, Wearing, and McLennan 1998; Readinger, Ross, and Crandall 2004; Thordsen 1991;

Wong, Sallis, and O'Hare 1997). This chapter describes how to conduct an effective CDM interview.

How can you go about "grabbing" the power that resides in the practitioner's experience? One way is to ask the practitioner how they do what they do, in a general or abstract way, as in "How do you predict tornados?" or "What's involved in doing the forecasting job?" This type of general question serves to divorce the practitioner's knowledge and skills from their lived experience.

In contrast, the CDM deliberately avoids generic questions of the kind, "Tell me everything you know about X," or "Can you describe your typical procedure?" Such generic questions haven't been very informative. One reason is that complex domains usually don't have simple, general, or typical procedures. Even if the work seems to be typical, we usually find many alternative types of action sequences even for routine tasks and situations. Furthermore, procedures change depending on style, the status of the equipment, and the skill level of the practitioner.

We came up with the idea of conducting the knowledge elicitation by asking people to tell us about previous incidents as a practical solution to a data collection problem. In a study of firefighters (Klein, Calderwood, and Clinton-Cirocco 1986), we had the notion of "shadowing" the firefighters, riding with them to fires, and interviewing them at the scene of the fire as the incident unfolded. We wanted to stand side by side with the firefighters and get them to "think aloud" (perhaps prompted by a few questions). But what seemed like a great way to get field data turned out to have a major glitch: Firefighting is an "on call" occupation, and there were stretches when there weren't that many calls. We soon realized we were likely to spend a lot of time (and money) waiting around for fires to happen so we could collect data, and that we weren't going to get very far by simply relying on observations. Instead, we used the downtime to collect firefighters' stories about some of their past experiences, using an adaptation of Flanagan's (1954) Critical Incident Technique. The retrospective method—asking people to tell us about previous incidents—arose out of necessity.

These CDM interviews rely on retrospection. In conducting CDM interviews we have to face the possibility of memory loss and distortion when significant time has passed since a to-be-recalled incident occurred. That is one of several reasons for probing non-routine, challenging events. By their nature, challenging events are going to call for whatever expertise a person can bring to bear on the situation. They evoke focused attention and depend on full use of skills. The outcomes often have more riding on them. For all these reasons, they are more vividly recalled than routine events.

This chapter covers two major topics: First, we describe the steps of the CDM—a method for mining people's real, lived experience and getting inside their heads to

understand incidents from their perspective. As we walk through the CDM interview process, we take you inside the interview to let you see an experienced CDM interviewer at work. Second, we discuss boundary conditions—when incident-based methods are most useful and when they are less so. We also describe variations of the CDM and how to adapt the method to different settings.

The Critical Decision Method (CDM) Procedure

The CDM is an intensive interview that often takes as long as two hours. In some cases such as weather forecasting where incident memories can be very rich, the CDM steps can even be broken up and conducted over several sessions. The CDM interview is conducted by two researchers. One interviewer acts as the primary facilitator, but also takes notes. The second interviewer is primarily responsible for taking a good set of notes and keeping track of the overall plan for the interview.

After making introductions, gathering some demographic information, and spending a few minutes establishing rapport, the main portion of the interview is carried out by making several "sweeps" through an incident. Each sweep constitutes a pass through the incident and builds on the previous sweep(s). Each is focused on eliciting specific types of information. At the end of the interview, the research team has a thorough understanding of the incident from the perspective of the interviewee.

In a CDM interview, the researcher tries to elicit information about cognitive functions such as decision making and planning and sensemaking within a specific challenging incident. The overall data collection strategy is to gradually deepen on critical cognitive points by making multiple passes through the incident. The research team has to get the story of the specific event and understand the cognitive demands of the task and setting.

The interview is conducted in four phases, or sweeps: (1) Incident Identification, (2) Timeline Verification, (3) Deepening, and (4) "What If" Queries (see figure 5.1). Each sweep uses different kinds of probes and perspectives and helps the participant recall events in greater detail.

In the following sections we provide a description of each sweep followed by a description of that portion of the interview from the researcher's perspective.

Sweep 1: Selecting an Incident

The initial CDM step is focused on identifying candidate incidents and selecting an appropriate incident for deepening. The precise type of incident will depend on the
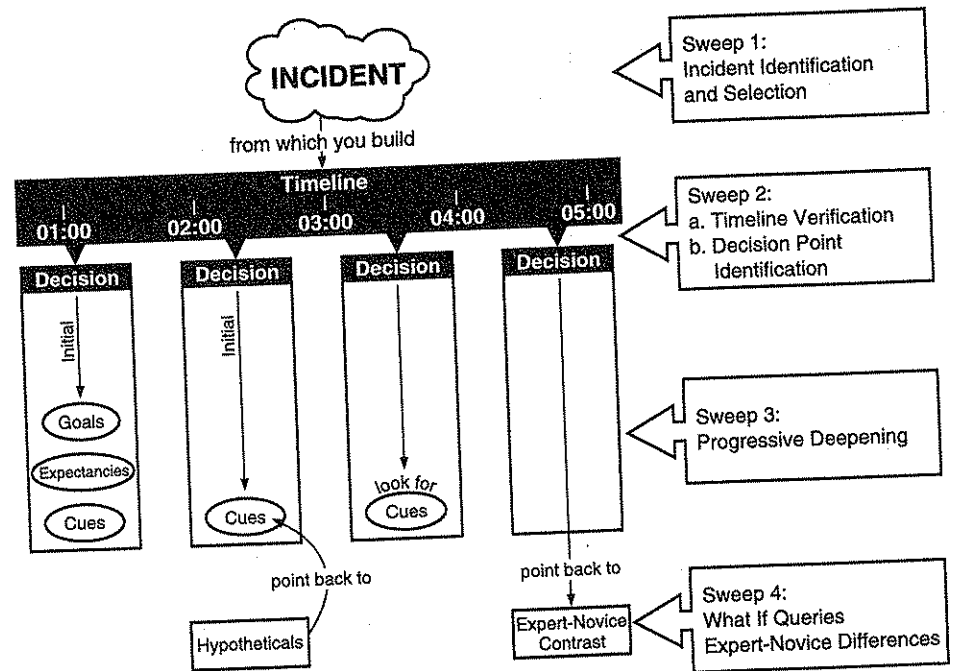
Figure 5.1
The structure of the Critical Decision Method (CDM) procedure.

nature of the project and goals for data collection. Traditionally, CDM has been used to examine nonroutine, challenging events, because these tough cases have the greatest potential for uncovering elements of expertise and related cognitive phenomena. For example, the interviewer might ask the participant to recall a time when his or her skills were particularly challenged, or when knowledge and experience really made a difference in the way the incident turned out. Generally, in this sweep you want to identify an incident that will contain cognitive components that go beyond background and routine procedural knowledge of the domain and that will enable you to learn about those components that characterize skilled performance and expertise.

Once you identify a good candidate incident, ask the person you are interviewing to provide a brief account of the story, from beginning to end. The initial account, and the content of the story, is the foundation for the rest of the interview. Table 5.1 offers a look at sweep 1 from the interviewer's perspective.

Table 5.1

Interviewer's perspective: What you're doing at this point in the interview—Sweep 1

Sweep 1: Incident Selection
Identifying a good incident has a number of elements:

- You are listening for a couple of key indicators of relevance. First, did the person you are interviewing have a role as a "doer/decision maker" in this event? Given the particular type of skill you are interested in, was this person in that role? For example, for a study of fireground commanders, the seasoned firefighter who was first on the scene and held command until backup arrived may have an interesting incident to tell. The firefighter who *witnessed* an incident involving exceptional command skill can't help you if he or she was not in the command role.

- What about telling participants ahead of the interview session that you are going to ask them for incidents? On the face of it, this seems like a way to save time for everyone. But it is risky. When they have advance notice, people mull over incidents. They are likely to rehearse a bit and in doing so they may begin to alter the story. They reorder events so that they will "make more sense." They discard the parts that seem not to "fit" or seem irrelevant. They smooth out all the edges, and leave out the embarrassing part where they made a mistake. These are exactly the details we want and need—so it's better to surprise them and get the story fresh.

- In asking for an incident, if you stress the unusual you are likely to get exactly that: weird stuff. Or you may get "critical decisions" but not of the type you are interested in. For example, a critical parameter for firefighters is whether or not life was lost. In the original firefighter studies, stories about those incidents were dramatic, often tragic, but they did not necessarily produce the kind of decision elements we were looking for. The issue is whether the person's decision making (or other cognitive event) had a direct impact on the outcome. If it did not, then the incident is probably not a good one for our purposes. (It is for this reason that the CDM was given its name.)

- When you are working in a new domain, you may find yourself wrestling a bit with whether an incident is worthwhile or not. You may need a few interviews before you have a better feel for what sorts of incidents the term "critical" is likely to elicit. There may be false starts and a need to use alternative opening queries.

- Be willing to sit quietly and let the person you are interviewing think about your question, even struggle a bit to come up with an incident. Do not rush them. If they say, "I can't think of anything," you might reply, "Let me say again what we are after." Repeat what you said before and add a bit of description, or rephrase your opening query in a slightly different way, and again give them time to produce something for you. Sitting in silence can be very hard, but the ability to tolerate silence is a key interviewing skill.

- Whatever your criteria, most people are going to have only a handful of incidents that fit it well. They won't need to sift endlessly through their whole past. In choosing an event to talk about, they will say, "Well, there was this one time..." They will give you an overview, an outline. If it doesn't sound worthwhile, you might say, "That sounds interesting, but we are looking more for incidents that..." and restate your criteria with some rephrasing. "Could you think of one that has more of that flavor to it?" Don't screen all possible entries. Once you hit one that sounds good, go with it. If it is the first one the person brings up, that's fine.

- What the person tells you gives you the content of the story. How they tell you the incident gives you the "bones," the basic structure, for the entire interview. In addition to the content, they have given you a sequence, organized into a series of segments. These incident accounts come to you initially as spoken stories that have an inherent structure and rhythm. Rhythm is about pauses—where there is silence in relation to where there is sound. Listen for the pauses, for where the person's voice falls for a moment before the next piece. Listen for the turning points, when the action or the entire scene changes. Listen for the words: "So then,..." These

Table 5.1
(continued)

are meaningful demarcations in the event. The sequence, the segments, and the pauses give you the frame of the story, dividing it into meaningful parts.

- You may need to move the person along in order to get through the initial account. People are usually eager to help, but they don't know exactly what sort of information you want, or at what level of detail. As they get into telling the story, they may dive down into the weeds, or they may wander off on a tangent and begin instructing you about standard operating procedures and general principles. You can help keep them on track by saying, "We're very interested in that, and I'd like to talk more about it in a bit; for now, can you give me a quick overview of this particular incident, so I have a sense of what happened from beginning to end?"
- In providing their view of the incident, the person defines the beginning and the end of the story. It can be informative to prove the beginning and end they provide. We often wonder, and sometimes ask, "What was happening right before this?" Or, "How did this turn out eventually?" Sometimes, what happened just prior to the person's starting point contains critical information for understanding the event itself. Sometimes the story has a second ending that provides a whole new perspective on the incident and the participant's role in it.

Sweep 2: Constructing a Timeline

The second sweep is aimed at getting a clear, refined, and verified overview of the incident structure, identifying key events and segments. This is a key step, because that structure will provide a crucial framework for the remainder of the interview. In addition, the person being interviewed often begins to recall events in greater detail and more fully relives the event. If you were observing a CDM interview, you might have a hard time telling where sweep 1 ended and sweep 2 began. Once you have identified an incident that appears to fit your project goals, and you have the initial incident account, it is appropriate to start verifying the timeline.

During sweep 2 the interviewer works with the participant to expand the initial, brief account of the incident. As the interviewer diagrams the sequence of events the participant might notice that something is out of sequence or that an event is missing and offer corrections and additional details.

Figure 5.2 contains an example of a timeline developed during an interview with a fireground commander. Clearly, the incident depicted here was challenging, and events were developing very quickly over a short time (approximately twenty-five minutes). Notice that the timeline is not laid out in equal intervals. The time hacks reflect timing of events as they actually occurred rather than fitting the incident into preset, regular time units.

In diagramming the timeline, the critical points (sometimes called "decision points") are when the practitioner experienced a major shift in his or her understanding of the situation or took some action that affected the events. They are the critical junctures in
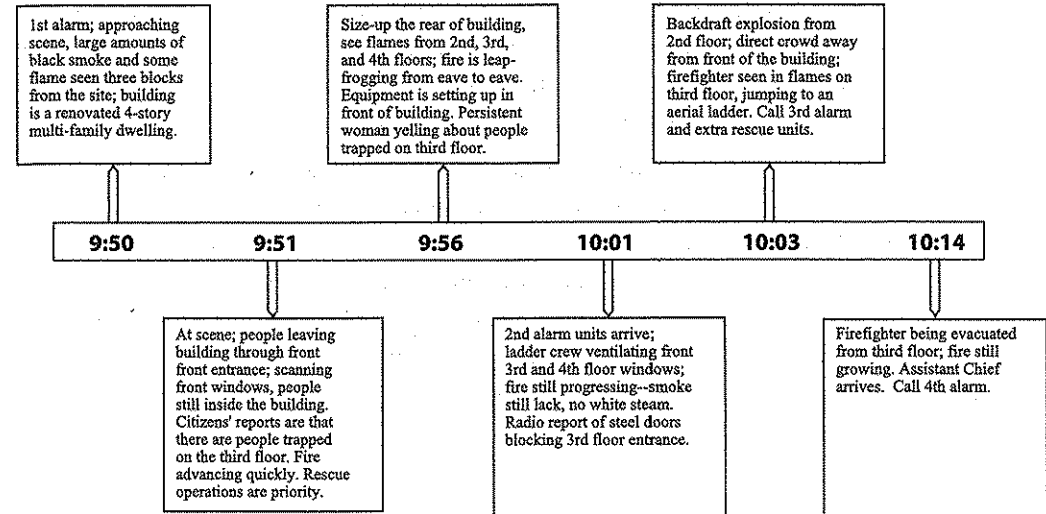
Figure 5.2
Example of a timeline from an interview with a fireground commander.

the event, turning points where the situation could have been understood or acted upon in several different ways (and not just the times when "good" decisions were made by the participant). Table 5.2 provides a look at sweep 2 from the interviewer's perspective.

As the sequence and duration of events, actions, perceptions, thoughts, and decisions emerge, the interviewers and participant arrive at a shared view of the facts of the case from the participant's perspective. Working through the incident in this way, interviewers are able to clear up inconsistencies, identify gaps, and fill in missing elements. Skilled interviewers are also able to begin tagging key segments and decision points to probe later on. As participants go back through their account, additional details emerge.

After interviewers have elicited and documented the incident and clarified and verified it with the participant, it is time to move to the next phase of the CDM interview: deepening.

Sweep 3: Deepening

This sweep is the most challenging, but also the most fun part of the interview. This is where you have the opportunity to get inside the expert's head and look at the world through his or her eyes. From the interviewer's perspective, the guiding question is:

What is the story behind this story? Based on the first two sweeps, I know what happened, who did what, and I know a bit about their role in the event. But what did they know, when did they know it, how did they know, and what did they do with what they knew? That's what Sweep 3 is designed to figure out.

Going beyond the time elements and basic facts of the incident, what were the participant's perceptions, expectations, goals, judgments, confusions, and uncertainties about the incident as it unfolded? What was he or she concerned about? What other options did he or she consider in making decisions? What information did he or she need and how did he or she get it? Critical Decision Method probes are used to deepen the understanding of the event—to build a comprehensive, detailed, and contextualized account of the incident from the decision maker's point of view.

Table 5.2
Interviewer's perspective: What you're doing at this point in the interview—Sweep 2

Sweep 2: Place the Incident on a Timeline
• A good way to start building a timeline is to go over the initial incident account, saying it back to the participant exactly as they have told it to you. Hold a mirror up for them. Ask, "Do I have the sequence and the details right so far?" Let them hear how the story sounded to you. They may realize something is out of sequence, and offer correction, and some additional details. Why it matters: the purpose here is for the two of you to agree on the overall incident. It's also the point at which the participant's memory begins to really engage. Inserting your language and your version of the story will muddy the waters. You want a clean, clear version of their incident as the basis for the interview, rather than your version of their version. Mirroring their account also shows that you are paying attention, and builds trust and rapport.
• What sort of timeline should you elicit? Do you need detailed and specific timing information? An approximate timeline? Are you interested in time (duration) or in timing (synchronicity, sequence of occurrence)? Perhaps it is enough to identify the sequence of events, because specific time designations do not have much meaning in a particular situation. In interview data elicited from NICU nurses (Crandall and Getchell-Reiter 1993), some incidents lasted less than three minutes, while others lasted several weeks. Interviewers had to decide in the moment what a useful time scale was going to be.
• Once you've settled on the right structure, the task is to overlay segments of the incident and key decision points on the timeline (or event line, or map). Typically we do this by creating a representation on a whiteboard or on 11 × 17 paper so there is plenty of room to write and work. You want the participant to be able to see what you are doing, what segments you are marking, and how you are labeling them. Engage him or her in the task by asking, "Do I have this right? About where on the timeline should we put _____?" The point here is not to make something that looks good (you can clean it up later if you like) but to generate an organizing framework that will help to keep you on track for the rest of the interview.
• You may have some pretty solid ideas about where the key decisions and situational shifts occurred. It is a good idea to get input from the interviewee at this point. You may think, "This occurred. It is a good idea to get input from the interviewee at this point. You may think, "This is a key decision," but he or she may say, "A guy right out of training would have known to do it that way, or he wasn't paying attention. That's standard operating procedure for this kind of situation." The decision clearly was important to the outcome, but it is not a critical decision in the sense that there was any other choice to be made.

Table 5.3
CDM "deepening" probe questions

Cues	What were you seeing, hearing, smelling, noticing etc.?
Information	What information did you use in making this decision or judgment? How and where did you get this information, and from whom? What did you do with the information?
Analogs	Were you reminded of any previous experience? What about that previous experience seemed relevant for this case?
Standard operating procedures	Does this case fit a standard or typical scenario? Is it a type of event you were trained to deal with?
Goals and priorities	What were your specific goals and objectives at the time? What was most important to accomplish at this point in the incident?
Options	What other courses of action were considered or were available to you? How was this option chosen or others rejected? Was there a rule that you were following in choosing this option?
Experience	What specific training or experience was necessary or helpful in making this decision?
Assessment	Suppose you were asked to describe the situation to someone else at this point. How would you summarize the situation?
Mental models	Did you imagine the possible consequences of this action? Did you create some sort of picture in your head? Did you imagine the events and how they would unfold?
Decision making	What let you know that this was the right thing to do at this point in the incident? How much time pressure was involved in making this decision? How long did it take to actually make this decision?
Guidance	Did you seek any guidance at this point in the incident? How did you know to trust the guidance you got?

Using the timeline and working from notes of the interview so far, the interviewer takes the participant back to the beginning of the incident and moves through it once again, taking the story one segment at a time. For each segment, the interviewer probes for additional detail and the participant is encouraged to elaborate on and deepen the incident account. During this sweep, interview probe questions are aimed at eliciting the cues and information available in the situation, the meaning they held for the participant, and the specific cognitive processes and functions they evoked. Table 5.3 contains the probes we have developed and routinely use for this sweep. Over the years, we have refined and revised the particular questions. Feel free to modify or add to this list as you discover useful questions and probes.

The interviewers will certainly not ask all of these questions for every key segment or critical point that is identified on the timeline, nor will they necessarily ask questions

about every topic contained in table 5.3. Knowing what probe to ask, when, and why, is a skill that develops with practice. Generally speaking:

- If a critical point on the timeline involves making an observation, the probes about information and cues are useful.
- If a critical point involves assessing or making sense of a situation or projecting a situation into the future, then probes about assessment and mental models are useful.
- If a critical point on the timeline involves making a decision, questions should be about decisions, obviously, but also about goals and options.
- At points where the story seems to refer to the participant's knowledge, then probes about the basis of choice and about experience are useful.

One of the interviewing skills that comes with practice is figuring out how to get good responses. Interviewers have to be ready to ask the same question in a variety of ways, because a probe that works well for one person may draw a complete blank from the next. When a probe doesn't elicit the information you expect it to, you have to know the reason for asking the question in the first place in order to come at the issue from another direction. Sometimes probing in this portion of the interview is like knocking at a closed door. You can knock once and walk away if the door remains closed; or you can knock again, maybe several times more, to see if you get a response. Table 5.4 offers a look at sweep 3 from the interviewer's perspective.

People sometimes evince physical or emotional reactions that suggest they are very much "in the moment." Firefighters begin to sweat. Pilots jump up, weaving their arms through the air and angling their hands to demonstrate a flight maneuver. People grab paper and pencil and start sketching the scene at a specific point in the incident to show movement of equipment and personnel. Weather forecasters draw simple charts and diagrams showing fronts and other features of weather situations. Sometimes people laugh at themselves, at how excited and involved they have become in telling us about the incident. Sometimes they choke up. Sometimes they cry. In these moments there is a profound sense of the participant reliving the incident and reporting on it as it unfolds. They are more "there," in that other place and time, than they are "here." Sitting in witness of this surging tide of memory can be an extraordinary experience—fascinating, intense, sometimes very moving. It can become difficult sometimes to keep writing, to stay in one's role as an interviewer and data collector. Sometimes the right thing to do is to put down your pencil and be willing to give your full attention to this person and his or her story, to share the recollection of a moment in time and space that had great meaning for this person. Eventually, the inter-

Table 5.4

Interviewer's perspective: What you're doing at this point in the interview—Sweep 3

Sweep 3: "Deepening" Using Cognitive Probes

- Your task at this point is to discover the story behind this story. Based on the interview so far, you know what happened and who did what. But what did they know, when did they know it, how did they know, and what did they do with what they knew? That's what you are there to figure out.
- Your specific questions and probes will depend in part on the goals of the project: what the key issues are, what aspects of expertise and/or the situation you identified as important in the preparation and framing activities. How you decide to explore the incident will also depend on what you heard in the initial account. What caught your ear? At what points did the SME say things like "We just knew ..." or "My gut told me that ..." or "It was obvious that ..." (As the interviewer, you may find yourself thinking, "It may be obvious to you, but it's not obvious to me, at least not yet.") The interview guide and the generic CDM probes are there to support active search and discovery, not to constrain or bound it. What's bothering you about the incident? What amazes you that you want to understand? Your own questions and curiosity are a critical compass for directing the interview.
- The incident account itself provides guidance and direction. Regardless of how the story has been segmented, take these chunks one at a time and work down through the layers of language and memory until you know everything about that part of the incident that the SME can tell you. Know it in its details and in what exists below its surface. The cognitive elements of this person's experience, of how she or he thought and acted inside the event, are under the surface of the story, waiting to be discovered.

viewee will take a deep breath and come back into the present, and you will pick up your pencil, and the interview will move on.

However, it doesn't happen like that every time. Even the most skilled and experienced of us have encountered interviews that yielded little, where participants are unable or unwilling to say what they know and to share their experience. It doesn't mean that the method is wrong, or that your skills are lacking. It just happens that way sometimes. Our suggestion is, if you have three interviews in a row that seem to get stuck, you should examine your interview strategy, your probe questions, and perhaps your choice of methodology (see the following section on boundary conditions).

At other times, participants can fly off in ten different directions, giving detail that is not relevant to the project. It is not unusual for participants to drift away from the specific case and shift into a generic discussion of how things are usually done or give a tutorial about the domain. One of the indicators that this is happening is a shift from first to third person pronouns (e.g., "You can always tell when things go wrong..."). One of your tasks is to keep the participant focused on the particular incident—the facts of the specific case. This doesn't mean you aren't interested in the participant's general knowledge—if it influenced this case, it is meaningful. But if they are launch-

ing into a tutorial on basic procedures, then it is time to shift the focus of the discussion back to the specific case. You can easily do this by asking, "Is that what happened in THIS case? Is that what you did THIS time?" It is your task to "steer" the interview, but it's important to do this with finesse and respect.

This sweep of the incident should yield a portrait of the participant's cognitive experience, skills, and knowledge. At the end of sweep 3, you should have a detailed, specific, and fairly complete picture of each segment of the event and of the overall incident.

Sweep 4: "What If" Queries

The final sweep of the CDM interview provides an opportunity to round out the interviewer's insight into the participant's experience, skill, and knowledge. Once again using the incident as a starting point, the interviewer poses various hypotheticals about the incident. These may be asked about the overall incident or about particular segments or aspects of the incident. One possibility is to invite the participant to speculate on how his or her responses in the event might have differed, or how the outcome might have been altered. Some suggestions for probes to use in sweep 4 are presented in table 5.5.

The "What if?" probes illuminate expert-novice differences and potential vulnerabilities for error in the domain. The probes allow you to expand the interview, using what actually happened as a springboard. Another way to expand the interview is by using props. Pictures, objects, photos/drawings, mockups, and storyboards all may be useful for depicting a hypothetical case or as a basis for posing several different hypothetical configurations. Alternatively, props can be useful for representing a concept that you want to get reactions to or ideas about. For example, we used props to enhance our questioning during a project examining women's concepts about osteoporosis. All the

Table 5.5
CDM probes for sweep 4

Expert-novice contrasts	If a novice had been in charge at this particular point in the incident, what type of error might she or he have made and why? Would they have noticed what you noticed? Would they have known to do X?
Hypotheticals	If [key feature] of the situation had been different, what impact would it have had on your decision/assessment/actions/plan?
Experience	What training might have offered an advantage in this situation?
Aids	What knowledge, information, or tools/technologies could have helped?

women in the study took nutritional supplements, including a calcium supplement. At a particular point in the interview we showed them an array of calcium supplements. Then, we showed them a bone (it was actually plastic, but it looked real). We said, "You've told us that calcium helps make your bones stronger. Can you describe for us how the calcium gets from here (the supplements) to here (the bone)?" Using these props was a much more effective way of eliciting their mental models of the role of calcium in bone health than simply posing an abstract question.

One of the key decisions that interviewers must make is how to allocate time for each of the sweeps and whether sweep 4 is essential to project goals. If an incident is very rich, two hours can seem barely adequate, and you may decide to forego most (or all) of sweep 4. There may be one question from sweep 4 that is essential to get to in the interview, and others that are considered extras that may be used if there is time. One of the CTA skills that develops with practice is the ability to think on your feet as the interview progresses, to figure out what information is most important and how to allocate the time you have.

We have found CDM interviews to be surprisingly intimate encounters. People share their experiences and sometimes gain new insights into what happened or realize aspects of their skills and knowledge that they may not have fully appreciated before. They may learn how they actually made critical decisions that they have been thinking about for years. One of the pleasures of doing CDM interviews is witnessing these moments of self-discovery. We have also found that it can be important not to end the interview session too abruptly. Instead, you should leave a few minutes to debrief, to answer any questions the participant might have, and to show your appreciation for the contribution he or she has made.

Boundaries and Limitations of the CDM

Although the CDM has some major advantages for doing knowledge elicitation, it is not always the best choice. There are situations, task domains, and project constraints where a standard CDM interview is not feasible or is unlikely to yield high quality data. We have encountered two types of conditions that limit the feasibility of the full CDM procedure as a method of CTA.

The first condition that limits our ability to do a full CDM is when there simply are no real experts, or even highly skilled practitioners, to be found. This can happen for a number of reasons. We have encountered domains in which there was only one real practicing "expert." In such cases, getting the person's time is practically impossible. Another possibility is that the job itself may be new, or has undergone a radical

transformation in technology and the way work is performed. Practitioners may not have had enough time to build skill in the domain. Or else the nature of the work somehow impedes the development of skilled performance. The domain may be one in which task feedback is difficult to discern, so practitioners are unsure when their actions have actually been successful.

If parts of the task are distributed across time, space, or personnel, then outcomes may become distant from the individual performer. It is very difficult for people to develop significant skill in the absence of clear, specific performance feedback. In a study we conducted several years ago of airport baggage screeners, for example, initial attempts to collect data using CDM were disappointing. Baggage screeners do stop bags that appear suspicious (and we came to understand the basis for their judgment, see Kaempf, Klinger, and Wolf 1994), but they get little if any feedback about how many they miss. And there are many opportunities to miss. On a typical shift in a busy airport, they were screening thousands of bags, often spending less than five seconds per bag. On a more practical level, the job has traditionally had very high turnover rates. As a result, many baggage screeners didn't develop a strong base of experience and robust skills.

A second condition that can limit the usefulness of CDM is one in which participants are unable to generate useful incidents. Combat-like conditions, where people work under severely stressful conditions and handle very high workloads, can create a blur of events that are difficult to recall as discrete cases. For a project on air campaign planning and targeting, we interviewed Air Force personnel who had been deployed to the Persian Gulf during Operation Desert Storm in 1990. Targeting personnel worked long shifts for many weeks prior to and during the conflict. They handled hundreds of targets, working from air campaign plans to acquire needed information for selected targets, briefing pilots, debriefing pilots, and conducting battle damage assessment. They were able to describe many aspects of the targeting task, but they found it extremely difficult to describe an intact case from beginning to end.

In the years since the CDM was first developed we have found that the same principles of incident-based probing can be used flexibly in a wide variety of types of CTA projects. In the sections that follow we describe some applications of CDM interviewing techniques that do not rely on retrospective accounts.

Adaptations of CDM

"Classic" CDM was developed to study decision making in naturalistic settings. The naturalistic decision making (NDM) perspective has widened its field of view, and

many NDM researchers, ourselves included, are investigating a range of cognitive phenomena that extend beyond decisions and decision making. (See the discussion of macrocognition in chapter 8.) In parallel, CTA practitioners have expanded, adapted, or altered the CDM technique in a variety of interesting and effective ways. These adaptations retain the emphasis on extracting incident-based data and reliance on people's lived experience as a basis for knowledge elicitation. They greatly expand the utility of the method. Most importantly, they demonstrate the breadth of possibility for application of incident-based knowledge elicitation.

CDM and Here-and-Now Incidents

Observation of experienced people at work is an important activity in CTA. In some disciplines where work is studied, such as the field of cognitive anthropology, observation is the main method. There are a number of ways in which observations and structured interviewing can be combined. In the study of baggage screeners in which CDM was impractical, the researchers developed an approach that relied on observations in airports—standing side by side with screeners and asking them questions about various aspects of what they were looking at and thinking about. Another example is from a project on the mediation of civil (e.g., noncriminal) legal cases (Crandall et al. 1996). The goal of the project was documentation of the cognitive skills involved in dispute resolution and how skilled mediators use their prior cases to help them plan for and carry out a mediation effort. The researchers were fortunate to gain the participation of a leading dispute resolution firm made up of attorneys and former judges. They allowed the researchers to shadow them during actual dispute resolutions.

Mediators typically conduct an initial mediation session with all parties present and then work with the disputing parties in separate rooms. The mediator shuttles back and forth between parties, discussing issues, listening to complaints and grievances, suggesting options, and (ideally) bringing the parties to agreement around a final settlement. Sessions often take several hours, some a full day or longer. In this project, the researchers stayed with the attorneys throughout dispute resolution, moving between rooms with them and eliciting responses to probe questions between the meetings with the individual parties. Sometimes these elicitation opportunities lasted only a few minutes. Sometimes they were as long as twenty minutes. Probes focused on the attorney's view of the mediation at that point in time. At the conclusion of the mediation, the researchers conducted CDM interviews, using the observations and the attorney's responses over the course of the resolution to fill out the attorney's incident account as the basis for additional data collection.

An abbreviated version of the CDM can be used to inform the researcher about a new domain and establish rapport with the participant. The method can also lead to the identification of potential leverage points, a tentative notion of practitioner styles, or tentative ideas about other aspects of cognitive work. The participant is asked to recall a salient recent case and describe his or her goals and activities. A timeline can be constructed, and, using that to anchor the discussion, the participant can be asked about any of a number of things, including information requirements (e.g., What information did you need or use to make this judgment?), mental modeling (e.g., As formation did you need or use to make this judgment?), mental modeling (e.g., As you went through the process of understanding this situation, how did you understand the problem scenario? Can you draw me a diagram of what it looked like?), and knowledge (e.g., How did this case relate to typical cases you have encountered? How did you use your knowledge of typical patterns?). This abbreviated CDM may be helpful when a full CDM interview is not feasible but the researcher wants to get a feel for some incidents.

CDM and Typical Incidents
Our discussion of the CDM brings with it the notion that a focus on critical decisions is often a good window into cognitive work. But not all CTA methods that are incident-based rely on the study of critical incidents. Some studies cannot rely on the study of critical incidents for the simple reason that not all real-world events comprise critical incidents. Thus, observations or interviews conducted during or immediately after real incidents will not necessarily end up speaking to critical decisions. And in some projects, the notion of "challenging event" doesn't make good sense at all, or it may not work well for the project goals.

Sometimes CTA researchers want to understand how things usually or typically work. In other cases, concerns around memory issues may lead the researcher to go after very recent events, to make sure memory for details is fresh. We have encountered these issues in some of the consumer projects we have conducted. If the product purchase or product use is a frequent event, we may ask for a typical experience of a particular type or we may request the most recent example. Similarly, in a project on physician-patient communication, we asked patients for examples from their most recent doctor visit, rather than a challenging one.

Variations on Use of a Timeline
In some instances, the requirement to elicit a timeline and structure the interview around it simply gets in the way (Militello et al. 2002). Interviewers find it easier and

more effective to move directly from the initial incident account to the deepening phase, particularly when the story segments are clear. How do you know when a timeline matters and when it is a frill? The answer is often contained in the domain and the incident itself; a timeline matters when the outcome of the incident depends on time or timing.

An example of a job where time and duration are critical comes from a project on landing signal officers (LSOs), whose job it is to help pilots land planes on the decks of aircraft carriers (Thordsen 1998). Landing an aircraft on a carrier at sea is a difficult, dangerous task. The fact that the landing strip is in motion is only a part of the difficulty. In addition, there is a very small window of opportunity, when the aircraft and the ship are lined up and synchronized and chances for success are optimal. If the window is missed (or doesn't open at all), the pilot must go around again. The LSO has just about forty-five seconds to make the determination to permit landing or to wave a pilot off and require another approach. In order to understand the LSO incidents, the CTA had to yield representations of the task and the LSO's cognitive activities to the second.

In some domains or situations, the outcome may depend not on time but on the particular *sequence* in which events occur relative to other aspects of the situation. In other domains or situations, spatial/geographic elements matter more than time. In a CDM study conducted with Alaskan pilots, researchers used maps to anchor incident segments and decision points rather than an actual timeline (Holbook, personal communication 2002).

Sometimes both sequence and geography matter. This is the case in many military or tactical situations. Linking aspects of a recalled incident to time or distance elements is the best way to ground the story. Another feature to consider is the length and complexity of the incident. Incidents that span several hours or more usually have many decision points, situation shifts, and multiple players. Here, a timeline can be a valuable aid to keeping all the details straight and in proper sequence. What matters is to figure out the structuring mechanism that will best support management of the interview in terms of making sense of the incident, keeping the sequence and details straight, and unpacking the important cognitive elements.

Conducting CDM Over Multiple Sessions
We mentioned earlier that CDM sessions can elicit detailed stories that take time to tell, retell, and deepen. In some domains, and for some types of incidents, fitting the CDM procedures into the standard two or so hours simply does not work. In these

cases, it can be productive to divide the CDM into two or more sessions. One approach that works well is to conduct a first session that includes sweeps 1 and 2: incident identification and selection followed by timeline development and verification. The second session includes deepening and "what if" querying. In a three-way split, the first session includes incident identification and selection. During a break, the interviewers transcribe the notes and prepare for the second session. The second session includes recounting the incident and timeline development. Again in the break, interviewers transcribe the notes and prepare for the third session. The third session includes timeline verification and decision-point identification, deepening and "what if" queries. Using this approach, the CDM sessions can be conducted over several days, enabling interviewers to document and absorb details of the complex incident and allowing participants to come to sessions refreshed.

The Knowledge Audit as Incident-Based CTA

The most thoroughly tested and validated adaptation of the CDM concept is the Knowledge Audit method (Hutton and Militello 1996; Hutton, Militello, and Miller 1997; Klein and Militello 2004; Militello and Hutton 1998). The Knowledge Audit examines the nature of the expertise needed to perform work skillfully. It structures an interview around a set of probes covering different aspects of expertise.

The CDM and the Knowledge Audit have sometimes been presented as a contrasting set: complex versus simple, incident-based versus general knowledge, and depth versus breadth of information. Hoffman, Coffey, and Ford (2000) regard the Knowledge Audit as a shortened or truncated CDM. In fact, the two procedures do share points of commonality, but offer distinct views of cognitive phenomena by using different elicitation techniques. The Knowledge Audit poses questions about specific cognitive elements that are characteristic of experts, based on the extensive research literature about expertise (e.g., Chi, Feltovich, and Glaser 1981; Ericsson and Smith 1991; Klein and Hoffman 1993). An example is the item designed to elicit information about perceptual discriminations:

Experts are able to detect cues and see meaningful patterns that less-experienced personnel may miss altogether. Have you had experiences where part of a situation just "popped" out at you, where you noticed things going on that others didn't catch? What is an example?

The Knowledge Audit was developed as a streamlined interview technique, designed for ease of use and accessibility. It is well suited to researchers who are new CTA practitioners. It can also be useful as the very first interview in a project because of the breadth of view it can provide.

The Knowledge Audit covers eight dimensions of expertise:

1. Past and future
2. Big picture
3. Noticing
4. Job smarts
5. Improvising/spotting opportunities
6. Self monitoring
7. Anomalies
8. Equipment difficulties

The purpose of the Knowledge Audit is not to demonstrate the importance of these factors—that is taken as a given. Rather, the purpose is to identify specific skills and perceptible patterns in the context of situations in which they have occurred and the expert's specific strategies for dealing with those situations. The Knowledge Audit is, therefore, useful in the exploration of apprentice-proficient-expert differences. Like the CDM, the Knowledge Audit draws on recall and description of examples. However, it bypasses the CDM requirement to identify and elicit a particular type of critical incident. Instead, the Knowledge Audit provides a structured interview format and set of predefined dimensions for eliciting and collecting examples. Knowledge Audit interviews produce a set of brief stories or minicases that illustrate how expertise plays a role in the particular domain. In a fairly limited time span and with a handful of participants, it is possible to generate a large set of examples organized around a well-defined and systematically applied set of dimensions. Working from the original concept, Knowledge Audits have been developed for use in studying cognitive aspects of team performance (Klein et al. 1999; Militello et al. 1999; Militello et al. 1994), macrocognition (Klein, Ross et al. 2003), and sensemaking (Klein et al. 2002; Klein, Phillips et al. 2003).

Incident-Based CTA with Teams

Many work situations and task functions are carried out by teams.[2] One can gain a very different picture of a work domain by examining the cognitive processes that underlie a team's skilled performance of tasks. Teams process information, make decisions, develop (and lose) situation understanding, detect and solve problems, and make plans (Cooke et al. 2000; Endsley and Jones 2001; Klinger and Thordsen 1998; Salas et al. 1995).

There are a number of team CTA methods currently in use (Klein 1998), and they include incident-based techniques. A version of CDM that has been adapted for use with

teams allows the researcher to elicit critical cognitive elements from multiple perspectives on a shared event. A CDM session is conducted with individual team members, who are each asked to describe an incident that is identified for them by the interviewer. The interview proceeds through the four sweeps. The cognitive probes may include questions regarding information sources and targets and aspects of coordination. The outcome is a data set of timelines, cues, goals, expectancies, and information sources, all gathered on the same incident from a variety of perspectives, roles, and functions.

Another adaptation is the team Knowledge Audit. It can be used to elicit aspects of team members' knowledge and skill regarding a specific task or set of tasks, examples, and events in which those skills were required. The team Knowledge Audit probes include: identification of decision makers; mission statement; developing and maintaining the big picture; information management; exposing expertise; team self-monitoring; and adaptability.

Summary

In this chapter, we described one method for using incidents to extract cognitive elements—the Critical Decision Method. We described the procedures for conducting a CDM interview and offered an interviewer's perspective on each of the CDM components. We examined the boundary conditions under which CDM is less likely to be effective, and we described some of the variations and adaptations that have developed to take advantage of the data collection opportunities that real, lived experience offers.