

te technical note techn

Review and Evaluation of Applied Research Techniques for Documenting Cognitive Processes in Air Traffic Control

January 1994

DOT/FAA/CT-TN93/39

This document is available to the public
through the National Technical Information
Service, Springfield, Virginia 22161



U.S. Department of Transportation
Federal Aviation Administration

Technical Center
Atlantic City International Airport, N.J. 08405

EXECUTIVE SUMMARY

The purpose of this report is to review and evaluate available research methods from cognitive psychology and human factors for the exploration and mapping of expert knowledge. A suite of appropriate methods is needed to investigate controller decision-making.

The first section of the report describes the purposes, research objectives, and scope of the study. Evaluation criteria are listed and used to assess each research technique. Research reports and the general literature are reviewed, and techniques for mapping cognitive structures and studying decision-making strategies are discussed in detail.

Previous work on cognitive structures is assimilated into Air Traffic Control (ATC) specific studies. ATC approaches include literature-based analysis; unstructured group discussion, protocol techniques, enhanced video recordings, cognitive task analysis, situation assessment, statistical techniques, and memory studies. General studies and techniques consist of scaling methods, dual-coding tasks, logical categorization, verbal-cuing tasks, and discourse-comprehension techniques.

Decision-making strategies research methods include withholding information (in an ATC setting), critical incident technique, critical decision method, and the use of verbal reports.

Each review of a method to explore cognitive structures or decision making is followed by an evaluation of its potential for problematic disruption of ATC operations, history of success in other environments, provision of meaningful indices of ATC cognitive processes, efficiency of data analysis, and potential for providing direct evidence of cognitive structures and strategies. A table summarizing each technique's ratings on these five indices is supplied.

Finally, an analysis and summary of the methods is offered and it is recommended that retrospective verbalization: structured probe interviews, video analysis, and statistical techniques (Pathfinder and multidimensional scaling (MDS)), be considered for practical application in an ATC-related study.

A comprehensive list of references is found at the end of the report.

1. INTRODUCTION.

The purpose of this task is to document the methods for examining the cognitive foundations of air traffic control (ATC). The focus is on two components of cognitive processing in ATC: cognitive structures and decision-making strategies. The cognitive structures of particular interest are those representing controllers' knowledge of ATC goals and principles; how to manage air traffic, the ATC/aviation environment, and when specific control actions are appropriate. The strategies of particular interest are the goal-directed methods used by controllers to manage traffic in specific time-bound situations. Because cognitive structures and decision-making strategies are not readily observable, it is necessary to identify methods and techniques that will produce valid, reliable behavioral data on which to base inferences about internal processes.

The results of two major activities are represented:

a. Selection and refinement of techniques for documenting cognitive process in ATC.

b. Measurement of controllers' cognitive structures and decision-making strategies.

1.1 PURPOSE.

This report describes the scientific precedents for the procedure proposed for use in documenting controllers' cognitive structures and decision-making strategies.

1.2 RESEARCH OBJECTIVE.

The primary objective, previously mentioned in paragraph 1.a., was to identify existing applied research techniques for documenting cognitive structures and decision-making strategies in ATC.

1.3 SCOPE.

A broad range of research techniques was reviewed (e.g., cognitive network analysis, concurrent verbal protocols, and retrospective interview techniques). Attention was focused specifically on the application of such techniques to ATC operations and other complex domains.

1.4 EVALUATION CRITERIA

In order for a specific technique to qualify as a candidate method for ATC, it had to meet several criteria:

- a. It must not disrupt ATC operations.
- b. It must have demonstrated success in other complex environments.
- c. It must provide meaningful indices of ATC cognitive processes.
- d. It must allow efficiency of data analysis.
- e. It must provide direct evidence of structures and strategies.

In section 2, an analysis of the techniques reviewed includes a tabular summary indicating whether particular methods and measures satisfy the various criteria.

2. SCIENTIFIC PRECEDENTS.

The literature review identified various techniques that have been used previously to investigate cognitive structures and decision-making strategies in complex environments. Only some of this previous research has been conducted within an ATC context. A larger proportion addresses other work environments such as fighter aircraft cockpits, industrial control rooms, and urban firefighting command centers. The review of previous studies provided a basis for selecting appropriate methods for the investigation of cognitive structures and strategies in the ATC domain. The following discussion presents a review of key techniques found in the literature reviews.

2.1 COGNITIVE STRUCTURES.

Cognitive structures refers to the content and organization of knowledge about ATC procedures; sector geography, flight characteristics, and previously experienced air traffic situations. These knowledge structures are thought to guide decision-making strategies, thus, they are key candidates for an investigation of controllers' cognitive processes. The remainder of this section presents highlights of studies that have addressed human cognition in the ATC domain.

2.1.1 ATC-Specific Studies.

It is widely believed that controllers build and maintain a mental picture of the current traffic configuration. The nature of this picture and the ways that its contents can be manipulated have been the subject of several studies. It is assumed that this current mental picture is just one of a family of structures that plays a significant role in a controller's choice of a strategy. Highlights of studies that investigated controllers' cognitive structures are reviewed below. Key methods and results are summarized.

2.1.1.1 Literature-Based Analysis.

Durso, Gronlund, Lewandowsky, & Gettys (1991), at the University of Oklahoma, conducted a literature review on cognitive processes in ATC as part of their investigation of the potential cognitive effects of automating flight data information. Using an ATC scenario, they discussed the cognitive processes involving short-term/working memory, attention, and generic knowledge that are likely to be affected by the replacement of paper flight progress strips with automated Flight Data Entry Notations (FDENs). Their focus on the dynamics of attention and memory enabled them to hypothesize that memory structures appeared to be based on the concept of chunking, or the grouping together of related aircraft in memory.

Working through the ATC scenario, they associated controller actions and decisions with psychological constructs such as prospective memory, concurrent processing, and implicit memory. Citing empirical research, they inferred that controllers develop cognitive structures for aircraft and airspace. They suggested, for example, that "controllers [may] tend to organize the airspace according to conflicts and potential conflicts" (p. 66). They made a case for the importance of paper strips in creating cognitive structures based on conflicts and conflict avoidance, which are described as "the most central goals of the controller."

Based on previous research by Rumelhart and Norman (1988), the Oklahoma researchers inferred the existence of controllers' knowledge structures for "facts...representations of overlearned surroundings...familiar event sequences...information about often experienced situations...successful strategies, and general rules" (p. 69). These researchers were concerned with the structure of the controller's pictures of the current situation and with the effects of changing the nature of input upon which the picture is based. They were also concerned with strategic planning and the effects of increased automation on this crucial activity. They raised numerous questions for operational test and evaluation of automated flight data and provided a useful list of references.

The method of literature-based analysis met only one criterion: It was non-disruptive of ATC operations. Although the authors implied that their method has been used successfully in other complex domains, particularly in Europe, their European references were primarily to empirical studies, not to literature-based analyses of ATC scenarios. The method of interpreting controller actions in terms of psychological constructs is inventive, but it is largely descriptive and inferential; it does not provide direct (empirical) evidence of structures or strategies. Direct evidence is provided, however, by many of the sources reviewed.

The authors appropriately called for "empirical research...to investigate the precise cognitive demands imposed upon the human operator whenever new systems are introduced" (pp. 79-80). They commented that "concerns about [the new ATC system] that were raised in [their] document need to be corroborated by behavioral research" and that their "theoretical analysis is no substitute for carefully planned and executed experiments" (p. 80). Empirical research is a key component of their ongoing work and of our current program of investigations into ATC conceptual structures.

Since no data were analyzed in the Oklahoma report, it is not possible to assess one of our criteria, efficiency of data analysis. It is necessary to go to the original sources cited in order to evaluate efficiency of data analysis. Our review includes many of the sources cited by these researchers.

2.1.1.2 Unstructured Group Discussion.

Research on the air traffic controller's mental picture of the traffic situation includes work performed by Whitfield (1979). In this study, off-line discussions were held with ten air traffic controllers who were in training. Questions were sent to the participating controllers before the discussions were held. Individual discussions addressed eight questions. The questions were designed to elicit descriptions of variations in the picture under radar and procedural control; information on how the controller constructs the picture, factors involved in losing the picture, traffic- and sector-related characteristics that cause difficulty in keeping the picture, reactions to and suggestions for research techniques, and reports of controller efforts to recall specific data.

These discussions surfaced opinions on sources of ATC workload, comments on differences between the comprehension abilities of experienced and less-experienced controllers, and references to a sense of foreground and background in the picture. Some comments suggested use of the picture as a plan against which to compare the actual behavior of aircraft. Predictive aspects of the picture were identified. Analysis of participants' comments appears to have been based on techniques of rational, but subjective, categorization.

In the discussion of research techniques to test the recall of aircraft information, some of Whitfield's controllers questioned the validity of testing conducted in a simulation environment.

Studying controllers' recall after coming off of shift from an active en route sector was considered more valid. Other suggested approaches included:

- a. Present a set of flight strips to a trainee and ask:
 1. Where are the aircraft?
 2. What is going to happen?
 3. How is conflict resolution to be achieved?
- b. Present a dynamic radar display with conflicts.
 1. Remove the data blocks.
 2. Ask the controller to identify aircraft and conflicts.
- c. Test controllers' knowledge of what is happening in other parts of the sector.
- d. Add a problem to the traffic situation and observe the nature and speed of controllers' responses, if any.
- e. Test for retention of specific conflicts (e.g., location and time of conflict, time of resolution).

Whitfield's primary research technique was the use of questions that the participants received in advance of the discussions. This technique suited his purpose which was to explore the concept of the controller's picture. Since the discussions were only "loosely" guided by the questions (p. 19), this technique falls into the category of unstructured group interview. Although this technique does not disrupt ATC operations and can provide preliminary insights, it is not particularly effective in providing data in a form suited to statistical analysis.

Techniques introduced by Whitfield's controllers are more specific. Their suggested questions are typical of probes that might be used in a structured interview. Since we are interested in the relationship between sector complexity and cognitive performance, a capability to add problems to sectors of varying complexity would be useful. Saving flight strips and using them to aid recall could be effective within a structured debriefing interview or video playback situation. While we could expect to have some unstructured discussions with controllers and training personnel, other more structured techniques will be needed to collect useful data.

2.1.1.3 Protocol Techniques.

Whitfield and Jackson (1982) used three key techniques for defining the nature of the air traffic controller's mental picture of the traffic situation: 1) interviews with training controllers; 2) elicitation of concurrent verbal protocols from mock-relief controllers who described how they establish the mental picture prior to assuming control of a position; and 3) elicitation of retrospective verbal protocols from controllers

during replay of radar displays and verbal transactions recorded during real-time simulations. The third technique included use of annotated flight strips from the simulation sessions.

Benefits and limitations of each method were discussed. Whitfield and Jackson suggested that presenting an unfamiliar sector to controller participants and allowing them to develop their own control strategies during simulation sessions may help in determining the factors that contribute to the emergence of a strategy. The authors emphasized the role of individual differences in developing mental pictures and control strategies. The difficulty of describing the picture in verbal terms was noted.

As compared to unstructured interviews, Whitfield and Jackson's collection of concurrent and retrospective verbal protocols offers more useful data. Verbalizations produced concurrently by subjects while they are engaged in work activities provide a meaningful index to cognitive processes. Although concurrent verbalizations give direct knowledge of the subject's thought process and strategy, it would not be possible to record them in an ATC operational or training environment because of the unacceptable level of disruption.

An issue raised about concurrent verbal protocols is whether the verbalizations really represent the actual thought process or a filtered, transformed version of that process (Ericsson & Simon, 1984). Since it is believed that a large proportion of the thinking in problem solving and decision making is unconscious, the actual process may be inaccessible. Carefully designed "think aloud" procedures have been used successfully in many domains to capture sequential verbalizations of subjects in problem-solving situations. These verbalizations can be recorded and later analyzed to identify decision points, alternatives that were considered and rejected, errors of intention, and patterns in individual styles of problem solving. Concurrent verbalization in ATC must, however, be reserved for a laboratory environment.

Retrospective verbalization is non-disruptive but problematic as a meaningful source of reliable data on cognitive processes (Ericsson & Simon, 1984). The problem is the unreliability of recall. For this reason, it is especially important to provide aids to memory such as video/audio tapes and marked flight strips. If recall is reliably cued, retrospective verbalization can yield direct evidence of cognitive structures and decision-making strategies. Both concurrent and retrospective verbalization can generate lengthy transcripts that require extremely time-consuming data extraction. Automated tools, such as the Shell for Performing Verbal Protocol Analysis (SHAPA) (Sanderson, 1990) have been developed and used successfully to aid in the analysis of concurrent verbal protocols.

2.1.1.4 Enhanced Video Recordings.

A paper by Roske-Hofstrand (1989) reported on the use of combined video and eye-movement recordings, which permit access to internal, unobservable cognitive processes. The technique involved superimposing recorded eye-movement and eye-fixation traces onto the corresponding taped sequences of human-computer interaction. To make analyzing the tape more manageable, the trace-enhanced video was then edited to create discrete, stylized summaries of the recorded interactions. Techniques from research on text comprehension were used to identify the boundaries of taped episodes. These techniques included collecting retrospective verbal protocols immediately after a taping session, prompting subjects for goal-subgoal relationships, and statistical scaling of eye-movement data. An ATC application of the enhanced video method yielded information on controllers' organization of aircraft traffic patterns. The author concluded that enhanced video records serve two purposes: 1) to provide cues for subjects in explaining their recorded interactions and their job tasks, and 2) to serve as archival databases for "naturally occurring [cognitive] behavior" (p.76). Individual subjects and groups of subjects can be compared on the basis of the empirically-derived, enhanced video episodes.

The techniques used by Roske-Hofstrand represent an effort to overcome the inherent imprecision of retrospective and video playback techniques. The objective was to make the taped record as scientifically valid and accessible as possible.

2.1.1.5 Cognitive Task Analysis.

2.1.1.5.1 Human Resources Research Organization (HumRRO) Study

In an ambitious attempt to derive instructional objectives for ATC training, researchers from HumRRO (Means et al., 1988) used techniques of cognitive task analysis to document the mental representations and strategies of expert controllers. Data were collected from three participating controllers, with experience ranging from 5-15 years, at the Jacksonville Air Route Traffic Control Center (ARTCC) training simulation facility.

The HumRRO researchers used a wide range of methods, including audio- and video-taping of real-time ATC simulations, post-simulation recall tasks, and concurrent/retrospective verbal protocols. All verbal interchanges were audio-taped. Concurrent verbal protocols were taped while the controllers were studying flight strips prior to simulation sessions. Recall tasks included sketching incidents and supplying flight strip information after a call-sign prompt was given.

Retrospective verbal protocols were collected both from the subject controllers and from other controllers who had played

supporting roles during the simulation sessions. Specific techniques of data analysis were not reported in the HumRRO paper but results appear to be based on transcribed protocols, on absolute numbers of cases observed, on comparisons of percentages, and on comparisons of estimated "average time to conflict." The HumRRO authors commented on the massive amount of data generated by their study and the need to discard some data that were recorded manually in the absence of automatic data capture capabilities. They concluded that their approach provided information which could not otherwise be derived.

The cognitive task analysis performed by HumRRO demonstrated the use of several techniques in combination to yield convergent evidence. Information gained from one technique helped to validate and complement information gained from other techniques

With regard to the selection criteria, techniques of cognitive task analysis meet several of the criteria (e.g., non-disruption, endorsement, preferred index). Traffic drawing tasks provide direct access to cognitive structures. Flight strip recall tasks allow insight into control strategies. Although verbal protocols taped during flight strip review are referred to as "concurrent" protocols, they were actually preview protocols that did not disrupt controllers' tasks during scenario execution in the simulation sessions. Similarly, post-scenario taping of relief briefings was non-disruptive of scenario execution.

2.1.1.5.2 Human Technology, Inc. (HTI) Study.

Researchers from HTI also conducted a cognitive task analysis to provide data for use in controller training (Redding, Cannon, Lierman, Ryder, Seamster & Purcell, 1990; Redding, Ryder, Seamster, Purcell & Cannon, 1991).

HTI's analysis focused on ATC expertise and individual differences in ATC skills, particularly in the area of goal/action prioritization. They used the following data collection procedures:

a. Unstructured Interviews. These discussions with supervisory Full Performance Level (FPL) personnel were conducted for purposes of orientation to ATC and initial investigation of mental models and prioritization strategies. Following preliminary data collection, 15 participants responded to a set of structured interview questions designed to elicit priorities for selection of strategies and techniques. The four most-experienced participants responded in writing.

Unstructured interviews meet only one of our criteria: non-disruption of ATC operations. They are useful as a source of preliminary insight and for establishing rapport with participants. They do not provide direct access to knowledge

structures or strategies. Structured interviews meet more of the criteria, if carefully designed probes are used.

b. Structured Interviews. Participants in three groups (expert, intermediate, and novice) were asked questions about their action priorities under normal and heavy workloads. They were asked to assign numerical priorities to nine actions on a scale from one (highest priority) to nine (lowest priority). Actions rated were:

- 1) scanning the plan view display (PVD)
- 2) sequencing traffic
- 3) calling and coordinating
- 4) determining crosspoints
- 5) accepting handoffs
- 6) data entry
- 7) reviewing flight strips
- 8) acknowledging pointouts
- 9) marking flight strips

Content analysis of the interview transcripts provided the basis for developing a procedural flowchart of ATC decision making with regard to aircraft separation. Ratings of action priorities were tabulated and graphed for visual comparison of the three groups under both workload conditions.

Since the focus of these interviews was on priorities, the technique may be described as "structured interview with focused probes." This technique is non-disruptive of ATC operations and provides self-reports on whatever is probed, in this case, priorities. This is a somewhat limited version of structured interviews with comprehensive probing, a technique endorsed by other researchers for providing a meaningful index of cognitive processes. HTI used more comprehensive probing in applying their critical incidents method.

c. Critical Incidents Interviews. HTI's critical incidents method, based on the method developed by Klein, Calderwood and MacGregor (1989) proceeded in the following steps:

1. Interview session: identify and describe unusual/difficult situations encountered by participants (five supervisor FPLs).

2. Second interview session: review of incident descriptions (situation, goals, sequence of actions); elicitation of possible alternatives to each action, with rationale for alternatives; use of probes about individual actions, goals/subgoals, and the overall situation.

This report discusses the critical incident technique and the critical decision method in detail in section 2.2.2. The

critical incident technique meets several of our criteria. It is non-disruptive; endorsed by other researchers; preferred as a meaningful index of cognitive processes; and amenable to efficient data extraction. It does not, however, provide direct access to knowledge structures or strategies.

Identification of decision points, and the use of probes, characterizes the critical decision method, which provides direct access to strategies. The probing method used by HTI exemplifies the critical decision method, which meets all but one of our criteria. It does not provide direct access to knowledge structures.

d. Paired Problem Solving. In HTI's method of paired problem solving, one participant sketched a difficult ATC problem on a sector map, constructed a problem scenario with drawings of several points-in-problem-development within the scenario, and generated a solution for the problem. In response to probing, the problem originator identified actions to be taken, hypotheses or goals associated with actions, perceived consequences of actions, information requirements for each action, and alternatives for each action. The problem was later presented to a second participant, whose task it was to provide a concurrent verbal protocol while developing a solution. The originator's solution was then read aloud and discussed.

In assessing the paired problem-solving technique against our criteria, we find it to be an innovative technique and a rich source of data but one that has not been used or endorsed by many researchers. Data extraction appears difficult. It yields self-reported information on strategies but little information on conceptual structures. It is non-disruptive of ATC operations.

e. Cognitive Style Assessments. HTI used two measures of controllers' cognitive styles. The Matching Familiar Figures Test was used as a measure of controllers' general tendencies toward "reflection" or "impulsivity" in decision making. The Group Embedded Figures Test was used as a measure of "field dependence" as contrasted with "field independence." The authors cited previous research to support linkages between cognitive styles and performance characteristics.

Assessment of cognitive styles is an attractive technique in that it offers a glimpse of what may be going on in controllers' minds. It is non-disruptive of ATC operations. It has been used and endorsed by many researchers. Measures of cognitive style, however, do not provide direct access to conceptual structures or strategies. Since cognitive styles may be correlates of structures or strategies accessed by more direct techniques, they may be helpful in explaining individual differences in structures and strategies. Data extraction is straightforward.

f. Dynamic Simulation (DYSIM). Use of the DYSIM facility provided data for HTI's performance modeling and an environment for structured problem solving. As input to performance modeling, five supervisors were classified as subject matter experts (SMEs). The SMEs completed four problem scenarios, consisting of two problem situations at two levels of complexity. Performance was video- and audio-recorded for playback and discussion. Computer printouts documented data entries made by SMEs and "ghost" pilots.

A week after completing the problem scenarios, SMEs were shown the video and asked a series of structured questions (probes) about their intentions, expectations, focus of attention, rationales for actions, importance of actions, goals, timing, concerns, and so forth. Their responses were used in constructing annotated scenario timelines for each participant. HTI's timelines can be understood as individual paths taken through the existing operations concept for en route controllers (Ammerman, et al., 1988).

The DYSIM provides a rich, high-fidelity environment for data collection. Although DYSIM problems lack the ultimate reality of live operations, controllers generally find such problems compelling and highly motivating. Data collection in the DYSIM is non-disruptive of ATC operations, although it must be scheduled with consideration for on-going training activities. It provides opportunities for audio- and video-recording that are not possible on the floor. Use of high-fidelity simulation techniques is widely endorsed by other researchers and advocated in the Federal Aviation Administration's (FAA's) National Plan for Aviation Human Factors (FAA, 1990).

Although DYSIM exercises do not provide direct evidence of cognitive structures, they can be used as a basis for eliciting self-reported strategies. Retrospective self-reports should be collected as soon as possible after the exercises are completed to minimize forgetting and guessing (Ericsson, et al., 1984) Verbal protocols about general intentions for managing expected traffic could be collected before DYSIM exercises begin (Schlager, Means & Roth, 1990).

Klein et al. (1989) recommended identification of critical incidents captured on the videotape. Collection and analysis of data on every recorded action is enormously time-consuming and may be of diminishing value beyond the critical incidents. Zachary, Ryder, & Zubritzky (1989) applied the COGNET modeling procedure to derive the panels for their ATC "Mental Model" from the detailed timelines.

If the purpose of the research is to gain direct access to conceptual structures and strategies, recorded DYSIM problems can

be used as a tool, but it may be helpful to use other data collection techniques before and after the problems are run.

In a second set of DYSIM activities, Zachary et al. (1989) used structured DYSIM problem solving "to capture performance strategies" (p. 13). Once again, the radar display and the participants' hand motions were recorded as they completed four different, low-complexity problem scenarios. Problems varied in amount of prioritizing required, level of planning and visualization required, time pressure, and completeness of flight data information.

Prompted by researchers, participants supplied concurrent verbal protocols while they were engaged in problem solving. Each scenario was stopped at pre-arranged five-minute intervals for discussion of the problem and participants' actions. HTI used standard techniques of protocol analysis as a basis for inferring strategies from participants' recorded verbalizations. Results of this analysis were used to derive a glossary of display strategies for planning and monitoring, control strategies, and workload-reduction strategies. The three groups were compared for their use of various strategies in the glossary.

As shown by HTI, the use of DYSIM affords a degree of control not possible in the live environment. For example, different participants can be given identical problems, thus providing a valid basis for within-group and between-group comparisons. Problems can be stopped for probing of the controllers' conscious thought processes. Variables such as problem difficulty/complexity, time pressure, and information availability can be held constant or manipulated.

2.1.1.6 Situation Assessment.

Lafon-Milon (1981) developed an empirical method to explore the spatial aspects of the air traffic controller's mental model. Expert and trainee controllers were shown eight static air traffic scenarios involving pairs of aircraft. One problem was a head-on approach, another was with overtaking aircraft, and six problems showed aircraft crossing tracks at 90 degrees. The controllers were asked to draw on paper the predicted relationship of the aircraft at the point of least separation. The resulting sketches were classified by whether they contained two-dimensional (plan or side view) or three-dimensional (plan and side view) information.

In a related report, Falzon (1982) critiqued current ATC display structures and emphasized the need for graphical representation of separation variables and their relationships. Citing the experiment performed by Lafon-Milon (1981), the author concluded that controllers represent vertical information in terms of relative values (i.e., separation between aircraft) not in terms

of absolute values such as those displayed in data blocks (e.g., altitude, speed). Falzon (1982) assumed that the air traffic controller constructs "a mental representation of the object of his activity" by filtering and coding relevant data in the light of operational goals (p. 297). He described the results, but not the methods, of previous research; only the Lafon-Milon study was discussed in detail.

The idea of asking subjects to represent the geometry of a future traffic situation might be used productively in a simulated environment. A conflict scenario could be stopped and the controller could be asked to draw what would happen without ATC intervention. The controller might also be asked to generate alternative interventions, describe these alternatives verbally, and draw the geometry of the future state that would occur for each alternative.

2.1.1.7 Statistical Techniques.

An interesting experiment which used multidimensional scaling (MDS) to explore the spatial form of the controller's pictures was done by Lapan (1985). This author developed "a methodology for direct and indirect reconstruction of cognitive maps (using MDS), as well as a diagnostic version of the methodology for studying mental rotation of three-dimensional objects" (p. 3). MDS is a scaling technique that uses direct similarity judgments as an input. These judgments are made by participants who use a rating scale to indicate the degree of similarity between two concepts and objects of interest. These similarity ratings indirectly reflect the participants knowledge about the stimuli (Kruskal & Wish, 1988).

Two groups of air traffic controllers were tested; group one was rated as highly qualified, while group two received only a satisfactory rating. Interobject distance judgments between targets on a radar display were processed using MDS. The conclusions drawn from the test results were that the skilled group perceived the targets as operating in three-dimensional space, while the less skilled group's space perception was only two dimensional.

Another research project which applied statistical methods to investigate the cognitive aspects of ATC was conducted by Landis, Silver, Jones, and Messick (1967). In their experiment, 51 air traffic controllers were classified into four groups based on job title. Within three of these groups, the controllers were further subdivided into levels of competence.

Thirty ATC problems were drawn on 8 X 10-inch map overlays of an air traffic sector. All possible pairs of stimuli were printed on 435 cards which the subjects sorted into 16 categories based on similarity. Using an early MDS procedure, it was determined

that the following dimensions were being used to judge the problems:

- a. amount of watching
- b. number of aircraft
- c. number of directions
- d. amount of analysis required
- e. severity of problem
- f. necessity for action

Mogford (1990a) summarized the work that has been done on the topic of mental models in ATC. He reviewed definitions of the term "mental model" and described the hypothetical uses of mental models in guiding behavior. Both conscious and unconscious aspects of ATC mental models were discussed. Mogford stated that the content of model partitions and the kinds of memory involved, as hypothesized by Whitfield (1979), needs to be corroborated by methods other than inference from direct verbal reports. Techniques such as MDS and cluster analysis appear to be effective methods for empirical exploration of ATC mental models.

MDS and other statistical techniques, such as cluster analysis, satisfy all but one of the selection criteria. They cannot provide direct knowledge of cognitive processes; however, MDS and similar techniques reduce the subjectivity of structural descriptions.

2.1.1.8 Air Traffic Controller Memory Studies.

Recall tasks have been used by several researchers to study ATC. Bisseret (1970) asked 3 groups of subjects (highly qualified controllers, average controllers, and trainees) to recall 7 aircraft attributes for sets of 5, 8, or 11 targets.

After a period of work with a DYSIM, Bisseret's subjects were asked to recall aircraft in three different ways. First, a sector map was provided on which they were to record as many aircraft as possible. Next, on a blank page, subjects were asked to recall as many aircraft attributes as they could from the simulation. Finally, the experimenter described the characteristics of new aircraft entering the situation and asked for conflict predictions.

Bisseret's results showed that the most highly qualified controllers had the best recall abilities. However, there was not much difference between average controllers and trainees. He found that there was a definite ranking of attributes across groups of subjects and workloads. Altitude and position information were remembered most frequently.

A recall task was also used by Sperandio (1974) to determine the effect of aircraft class membership (e.g., conflicting or non-

conflicting) on controller memory for the aircraft remaining under control at the moment of being relieved by another controller. Subjects filled out blank flight strips with the information they recalled (e.g., call sign, aircraft type, flight level, position). Following the recall task, subjects were asked to return to their workstations and identify aircraft that had been in conflict or had been targeted for a control action to resolve a conflict. Analysis of the data focused on the number of aircraft recalled and the kinds of characteristics recalled. Results are interpreted as confirming Bisseret's (1970) general hypothesis that memory is organized and structured by the operational task.

Mogford (1990b) completed a research project on the controller's picture. An experiment was conducted with air traffic controller trainees to determine if the accuracy of the picture (as measured by recall of aircraft data) is important for ATC. Subjects were given a picture assessment test, an aptitude test battery, and memory tests. Multiple regression was used to determine if the trainees' final scores in series of simulator-based exams could be predicted by the picture assessment and other test variables. A number of the variables emerged as predictors of radar simulation final test scores.

A project on air traffic controller memory was completed by Vingelis, Schaeffer, Stringer, Gromelski and Ahmed (1990). The first report reviewed presented the results of the first year of work on a 3-year study to "enhance National Airspace performance by developing a set of practical and effective memory aids to improve controller performance of tasks where memory is a critical element" (p. vii). The research was intended to develop an understanding of memory in controller performance, identify memory problem areas, describe potential memory aids, and evaluate candidate memory aids.

An FAA task force on ATC operational errors highlighted a number of contributory factors including controller memory lapses and faulty information scanning (FAA, 1987). Vingelis, et al. (1990) stated that an understanding of controller memory functioning is needed in order to prevent memory lapses and design effective memory aids to be incorporated into new ATC systems.

Vingelis, et al. (1990) reported on a review of the literature on working memory as related to ATC tasks. They suggested that human memory is composed of three subsystems: sensory storage, working memory, and long-term storage. Sensory storage and long-term memory were briefly mentioned. A discussion focusing on working memory and the issues of rehearsal, capacity, chunking, forgetting (decay, retroactive inhibition, proactive inhibition), attention, and cognitive resources was concluded.

The review of previous research on air traffic controller memory by Vingelis et al. (1990) revealed that the ability to memorize air traffic data increases with experience and decreases with the amount of traffic. They cited previous research to support the claim that ATC information is hierarchically organized in working memory. It also appeared that increased workload caused controllers to adjust their operational practices to allow optimum performance. This may have included processing fewer aircraft-related variables.

Vingelis et al. (1990) offered a definition of "controller tactical working memory [which] consists of functional requirements (attention and rehearsal), contents (aircraft data, position, etc.), capacity (7 +/- 2), and limitation (interference) within a 3- to 5- minute tactical window" (p. 17). They acknowledged the influence of controller training, ATC procedures, and personal preferences, which reside in long-term memory and guide search and retrieval mechanisms, but stated that the study would be oriented toward working memory. Surprisingly, Vingelis et al. (1990) did not cite research showing a more constrained capacity for working memory (i.e., four to five chunks).

A cognitive model, based on Rasmussen's work on hierarchical levels of behavior, was developed by Vingelis et al. (1990). Rasmussen's method of human error categorization was reviewed as related to skill, rule, and knowledge-based operational activities. Operational categories of ATC system error were then related to Rasmussen's error descriptions, the authors' cognitive model, and to possible memory factors which could be operative.

Vingelis et al. (1990) described a set of 16 memory aids, ranging from special flight progress strip modifications to automated decision aids. Each memory aid was related to the previously discussed operational error categories in terms of how it could reduce errors. The aids were evaluated by experienced and inexperienced controllers for face validity, usability, feasibility, effectiveness, cost, and testability. Further testing of the top ranking memory aids was recommended.

Bisseret's (1970) technique of asking controllers to recall air traffic information can provide data on memory efficiency and the organization (or prioritization) of this knowledge. It may be disruptive of ATC operations if memory probes are required at regular periods. However, recall could be requested after the end of a shift, as reported by Sperandio (1974). It has been used in at least one other complex environment (Fraker, 1989) and it meets the other three selection criteria.

The approach used by Vingelis et al. (1990) was primarily based on a literature review with ratings of the suggested memory aids

by air traffic controllers. It is non-disruptive of ATC operations and an approach endorsed or in common use in similar studies. No specific empirical research methods were described

Vingelis et al. (1990) provided some useful information about research on controller memory but focused on working memory as opposed to long-term memory. The cognitive structures which are of interest in the current study are resident in long-term memory. The authors appeared to be oriented toward developing memory aids as opposed to conducting empirical research on the structure and dynamics of memory in the ATC environment.

Recall techniques can be non-disruptive of ATC operations, are endorsed by other researchers, are a meaningful source of information, and may offer direct access to knowledge structures (especially with regard to assessing the relative importance of different kinds of knowledge).

2.1.2 General Studies/Approaches.

A great deal of research has been conducted to elucidate cognitive structures in domains other than ATC. Some studies use a direct approach to capturing mental models and other internal structures (i.e., subjects verbalize or graphically depict their mental models). Other techniques use sophisticated statistical procedures, such as Pathfinder (Schvaneveldt, Durso, Goldsmith, Breen & Cooke, 1985) and MDS (Kruskal et al., 1988), as a basis for determining the ways in which mental representations of objects and concepts are linked together. Highlights of selected general studies are presented in the following brief reviews.

2.1.2.1 Scaling Techniques.

A report by Schvaneveldt, Durso, and Dearholt (1985) described the Pathfinder algorithm for obtaining insight into the way people's knowledge is organized. The algorithm generates a family of link-weighted networks from any set of distance data. Distance data are usually similarity ratings, which imply the "psychological" distance between concepts and terms. A derived network consists of a set of concepts and links that directly connect pairs of concepts that are highly related. This represents the basic organization of existing knowledge structures or elements based on an individual's experience. This method has been successfully applied to determine the organization of pilots' knowledge as a basis for designing a menu system (Roske-Hofstrand & Paap, 1985), and also to distinguish the expertise of novice and expert fighter pilots (Bisseret, 1970).

Roske-Hofstrand et al. (1985) used the Pathfinder cognitive network analysis tool (Pfnnet). Network analysis was used to facilitate the design of an electronic menu system for cockpits

The rationale was that complex management systems need to be compatible with the pilot's cognitive structures. Pathfinder was used to derive a representation of how information in the menu system should be grouped together. The menu design that resulted from Pathfinder led to superior performance when compared to other menu organizations. Scaling methods such as Pfnnet and MDS have traditionally been used to quantify structures and the dimensions of human knowledge which are used to organize semantic elements (Schvaneveldt, Durso, Goldsmith, Breen, & Cooke, 1985). Converging evidence suggests that these techniques are scientifically valid and rigorous.

2.1.2.2 Dual-Coding Tasks.

In a chapter on the "Manipulation and Use of Representational Information," Paivio (1986) discussed research tasks that require subjects to compare perceived and imagined objects, to compare symbols, to make mental transformations, and to perform computations based on representational structures. These types of tasks are probably similar to those performed by air traffic controllers. Findings were interpreted in terms of dual coding theory. Paivio (1986) concluded that mental representations of imagined objects and of perceived objects "contain similar structural information ... that ... can be used and manipulated in similar ways" (p. 177). Evidence was presented for a relationship between subjects' high imaging ability and superior performance. This finding suggested that it might be worthwhile to group individuals according to their imaging ability when using them as subjects in studies.

Comparison tasks like those devised by Paivio might be useful in exploring how mental representations are constructed and manipulated. Structured interviews could include comparison tasks. Data extraction might be time-consuming, but the other measurement criteria would be met.

2.1.2.3 Logical Categorization.

Rasmussen (1979) described and illustrated various cognitive structures developed by operators in industrial workshops and control rooms. These mental representations of a complex system's structure and functional properties were derived from analysis of concurrent verbal protocols. They were categorized according to levels of Rasmussen's abstraction hierarchy: physical form, physical function, functional structure, abstract function, and functional meaning or purpose. Rasmussen (1979) defined a mental model "as the internal representation of the properties or constraints in the environment which determine the interrelations among the data which can be observed from the environment" (p. 9). He made a distinction between quantitative mental models and qualitative mental models. Strategies used by operators in coping with complexity were discussed within the

context of this approach to elucidating the form and structure of mental representations. This paper implied that ATC cognitive structures may include functional models, state pattern models, and implicit models such as procedural rules. Rasmussen's abstraction hierarchy provided one model for categorizing various ATC cognitive structures.

Rasmussen's methods were primarily logical and taxonomic in keeping with his objective of developing a standard taxonomy of mental models. Models found to be important in diagnostic tasks are characterized, formalized, and categorized from a functional perspective. A taxonomy of models is derived from analysis of concurrent verbal protocols. Taxonomic categories are physical form, physical function, functional structure, abstract function, and functional meaning or purpose. These categories occupied levels in Rasmussen's abstraction hierarchy. Use of this approach to categorizing ATC cognitive structures might be helpful in associating different categories of structures with different types of ATC situations. Validation of the categories could be attempted by application of statistical techniques such as MDS and Pathfinder.

If Rasmussen's categories were validated, use of his abstraction hierarchy would satisfy several of the selection criteria (i.e., non-disruption, endorsement, and meaningfulness). By categorizing cognitive structures and associating them reliably with specific ATC situations and specific control strategies, we would be developing a basis for predicting controller behavior.

2.1.2.4 Verbal-Cueing Tasks

An article by Johnson-Laird (1981) was a precursor to his book on mental models (Johnson-Laird, 1983). Both the article and the book investigated issues of the mental representation of verbally-cued information (e.g., how internal models of logical premises were represented mentally and how they are used in cognition). Methods employed to confirm the existence of mental models were drawn largely from the fields of logic and linguistics (e.g., logical inference, communication theory). Logical evidence was presented for the manipulation of mental models, for their dimensionality, and for their dynamic characteristics. The author stated that manipulation of mental models is a demanding cognitive activity, one that can exceed the limits of human capabilities. Underlying mental models may provide "prototypical information" about the objects of perception but not be accessible to consciousness as a mental image. Empirical evidence from drawing exercises is interpreted as supporting the hypothesis that a mental model can be constructed from verbal description but also draws on "general knowledge and other relevant representations in order to go beyond what is explicitly asserted" (Johnson-Laird, 1983, p. 180).

Implications for the study of ATC mental models are: 1) that the controller may build up internal models of verbally-presented information (e.g., communications, rules from the ATC handbook; aircraft position and weather from verbal pilot reports); 2) that mental models may underlie expectations for the future traffic situation, and 3) that the potential for "losing the picture" may be related to the level of mental manipulation required to "test" a mental model against actual traffic status. Exploration of the first implication involves linking specific sources of information with specific cognitive structures. A study to do that would provide important input to ATC training programs. Exploration of the third implication would require additional research beyond the scope of the current task.

Exploration of the second implication would be possible in the current task if verbal protocols could be collected while controllers were reviewing flight strips prior to the start of a DYSIM exercise. These would be preview verbal protocols in that they would precede any concurrent verbalizations during the simulation exercise itself. This technique has been referred to as "knowledge engineering" in other contexts.

A plausible hypothesis is that controllers prepare for their shifts on specific sectors by activating generic cognitive structures (i.e., schemas) for typical traffic patterns, given the time of year, the time of day, and values for other variables that can be instantiated on the basis of prior experience. The assumption is that case-based expectations are built up and represented as schemas for typical patterns that have been observed over time. Similarly, expectations are assumed to be associated with control strategies that have been successful in the past.

This assumption recalls an important point made by Rasmussen (1979, p. 8):

Human activity in a familiar environment will not be goal controlled, it will rather be oriented towards the goal and be controlled by a set of rules which have proven successful previously... The efficiency of humans in coping with complexity is largely due to the availability of a large repertoire of different internal models of the environment from which rules to control behavior can be generated ad hoc.

If this is true in ATC, then expectations for traffic will be linked to highly-practiced rules (i.e., procedures) for controlling the traffic. Expectations become the basis for organizing traffic, and successful experience in managing traffic becomes the basis for linking rules with expectations. Both the

expectations and the rules, as well as the links between them, must be represented in a cognitive structure of some kind.

It would, thus, seem useful to probe controllers' expectations before a control session with simulated or live traffic. The controllers' expectations for the traffic and for their strategies for managing it can later be compared to what actually occurred and how situations were handled. If traffic conformed to expectations, did the controller stick to the planned strategy? If traffic deviated from expectations, how quickly was the deviation noticed and what effect did it have on the control strategy? Have the controllers' expectations for traffic in this sector changed as a result of this particular control session? If expectations have changed, has there been an accompanying change in the generic strategy for this sector?

These research questions can be investigated to some extent by Johnson-Laird's techniques of logical inference and drawing exercises. These techniques meet several of the selection criteria (i.e., non-disruption, endorsement, and meaningful index). In addition, drawing exercises provide direct access to cognitive structures. Although Johnson-Laird (1983) used these techniques to study models of verbally-cued information, they could be used to study models of visually-cued and aurally-cued information. These techniques would, however, need to be combined with preview verbalization, observation, structured interviews/retrospective verbalization with probes, video/audio replay, and statistical analysis in order to gather convergent evidence for a cognitive network of expectations and strategies.

2.1.2.5 Discourse Comprehension Techniques.

Based on research in the field of discourse comprehension, Morrow and Greenspan (1989) focused on readers' development of situation models from textual descriptions and from mental elaboration through inferencing. The authors pointed out that a reader uses textual cues in determining the relevancy of elements of the described situation and that these cues essentially prime the reader's memory to make relevant information accessible for elaboration of the situation. Methods included presenting narrative text and asking subjects to locate items and characters mentioned in the text. This was somewhat analogous to asking controllers to draw pictures of sector situations.

By implication, this article raised the question of what determines the accessibility of ATC information in the controller's memory. Controllers are readers in the sense that they read the situation by integrating visual, verbal, and mental information. The article suggested that the reader's perspective and focus contribute to determining relevance and accessibility. If the goal of communication is to share a model of a particular set of situations, findings from this kind of work in discourse

comprehension may have strong implications for flightdeck-ATC communication models and for all ATC interchanges.

One implication is that pilots and controllers may read the situation differently given their differences in perspective and focus as well as the differences in the visual information available to each of them. Another is that mental inferencing may lead pilots and controllers to different conclusions given the same verbal information. A third implication is that training in reading situations may be as important for pilots and controllers as it is for readers of textual discourse. Readers who have not developed a sensitivity to textual cues will have problems in comprehension. Just as training can improve discourse comprehension, it should be possible to train pilots and controllers for improved sensitivity to visual and verbal cues. All of these implications suggest topics for further research.

Techniques from the field of discourse comprehension have not been used much, if at all, in investigation of controllers' cognitive structures. They meet several of the selection criteria (i.e., non-disruption, endorsement, meaningfulness, and direct access to knowledge structures). Some of these promising techniques could be incorporated into structured debriefing interviews. Controllers could be asked to read a description of sector traffic and draw it from memory. They could be asked to acquire the description visually or aurally and then draw it from memory. Several variations on this technique would be possible. Although it is beyond the scope of this current task, one variation could include pilots, as well as controllers, working with the same written, visual, and aural descriptions of traffic situations.

2.2 STRATEGIES.

A strategy is a goal-directed use of resources in response to a situation over time. Much has been written on methods for investigating human problem-solving and decision-making strategies. Only two studies were located that addressed controller strategies specifically. Key aspects of these approaches are reviewed below.

2.2.1 ATC-Specific Studies.

Most methods for identifying operator strategies involve observation. Variations on this method include passive observation, probing the operator at certain intervals, and withholding information (thereby requiring the operator to request it). The two studies reviewed below involved withholding information from the controllers. Observers recorded the type of information requested and the ordering of requests. When coupled with other data, such as performance and workload ratings, this

information provides convergent evidence of the nature of controllers' strategies.

2.2.1.1 Withholding Information

The main argument by Sperandio (1971) was that the choice of a strategy for controlling air traffic depends upon the characteristics of the operator, task, and workload. An interesting observation is that the particular strategy can also influence the level of workload experienced by the controller. Sperandio investigated approach controllers' strategies using a realistic radar display. The number of aircraft was varied, as well as whether they had been cleared to land. The only data provided to the controllers were the aircraft locations as displayed on the radar screen. Thus to sequence the traffic, controllers had to request information pertaining to flight levels, speeds, aircraft type, and so forth. In order to gain an insight into controllers' strategies, the order in which controllers requested information was recorded, as were routing solutions (e.g., direct approaches, semi-direct approaches, use of holding patterns, and separation between aircraft). Results showed that controller workload increased with the traffic but that the increase was regulated by the choice of a particular strategy.

A study by Coeterier (1971) also investigated controller strategies associated with the terminal environment. Radar pictures of the terminal area were used to understand how restrictions on maneuvering space affected variations in controller strategies. As in Sperandio's study, certain pieces of information were withheld from the controllers. The experimenter recorded the order in which information was requested. Results revealed that strategies vary with variations in flight progress of aircraft and existing maneuvering conditions in the airspace.

The technique of withholding information meets several of the selection criteria (i.e., endorsement, meaningfulness, efficiency of data extraction, and direct access to cognitive structures). Since it is disruptive of continuous task performance, however, this technique can be applied only in a laboratory setting.

2.2.2 General Studies/Approaches.

Elucidating strategies requires going beyond traditional measures, such as errors and latency of performance. The methods described below focus primarily on verbal reports as data. Verbal reports provide considerable insight into the cognitive processes employed in performing a task, especially where cognitive activities play a significant role. Four studies are reviewed that provide direct, meaningful insight into the present task. Additional articles are listed in the bibliography.

2.2.2.1 Critical Incident Technique.

In a pioneering article, Flanagan (1954) described the "critical incident technique" as a method for structuring the experimenter's observation of an operator performing a task. The critical incident technique is a set of procedures for collecting direct observations of human behavior in a way that facilitates their potential usefulness for developing psychological principles about problem solving. The first step involves classifying critical events, that is, incidents that are sufficiently complete to permit inferences and predictions to be made about human performance. The second step is to define the performer's aims, goals, and intents. The third step is to identify any planning and decision-making aspects that were part of the process. Flanagan noted that it is important that the observers have knowledge of the activity being observed. The critical incident technique is an ideal method for strategy description in that it provides a functional description of the activity in terms of specific behaviors. It is on this method that Klein, et al. (1989) base their critical decision method, as described next.

2.2.2.2 Critical Decision Method.

Klein et al. (1989) and his colleagues described a critical decision method for modeling tasks in naturalistic environments characterized by high time pressure, high information content, and changing conditions. This method builds upon Flanagan's critical incident technique. The method uses a set of cognitive probes or questions in a retrospective interview to determine the basis for conceptual discriminations, typicality judgments, and critical cues. The probes focus discussion about concrete events as opposed to general rules and procedures. Types of probes include "cues" for extracting the sources of information the operator uses, "knowledge" used during the situation, "goals" of the operator at specific times during the task, the "basis" for taking a certain action, and so on. The probes are tailored to specific events, thereby achieving a meaningful and specific level of detail. The method provides a rich source of data and has been successfully applied to complex, real environments such as urban and wildland firefighting, tank platoon commanding, and paramedics' responses to emergency situations.

Both the critical incident technique and the critical decision method include a set of techniques: structured observation; classification of critical events; logical inferencing; definition of aims, goals, and intents; and identification of planning and decision-making processes. The critical decision method adds the use of probes in a structured, retrospective interview. Together, these techniques meet all but two of the selection criteria. Although extraction of data can be time consuming, these techniques are well suited to elucidating ATC

strategies. They provide indirect access to cognitive structures.

2.2.2.3 Verbal Reports.

A book on verbal protocols by Ericsson et al. (1984) provided an extensive treatment of verbal reports as well as a comprehensive model of human information processing that distinguishes the various processes underlying verbalization. One of the main arguments was that the accuracy of verbal report data is highly dependent on the methods used for extracting information. When direct probing is used, that is, probes or questions that refer directly to the task or situation, verbal reports are likely to reflect accurately the cognitive processes underlying the task. When emphasis is placed on concrete cases, subjects are provided a context for recalling how they actually performed the task. When probes are indirect and general, on the other hand, subjects may be forced to draw on information that had nothing to do with the actual processes used.

Ericsson et al. (1984) identified the following two situations where they would recommend the use of verbal reports:

- a. Where information is of a verbal nature, verbalization is immediate.
- b. Where information is not of a verbal nature but verbal coding is easy, verbalization of the information would not introduce any notable distortion.

This technique is not well suited where verbal coding of information is difficult for the subject or where intermediate cognitive processes, such as abstraction and inference, are likely to interfere with the accuracy of verbal reports.

A distinction is also made between two types of verbal reports: concurrent and retrospective. Concurrent reports can be intrusive because the subject is required to "think aloud" while performing a task. This method is not recommended for tasks that involve communication, such as ATC. For these tasks, retrospective verbalization is the only feasible method.

An article by Leplat and Hoc (1981) focused primarily on retrospective verbalization. The authors agreed with Ericsson et al. (1984) that one of the main criticisms of retrospective techniques is their questionable validity. However, it is possible to define instructions that do not systematically orient subjects to reflect on and interpret their own cognitive processes. As did Ericsson and Simon, Leplat and Hoc argued that the type of probes used and the instructions given to the subject need to be carefully controlled. Another factor that can greatly influence the nature of retrospective verbal reports is the type

of language used by the experimenter. It is important that the analyst knows the lexicon used by the subject and possesses technical knowledge of the task. Because ATC operations cannot be disrupted for obvious reasons, reliance on retrospective verbal reports is necessary in attempting to identify cognitive structures and decision-making strategies in ATC. Every attempt must be made, however, to overcome the deficiencies of retrospective verbalization. Control of instructions and wording of probes is essential. Researchers must be familiar with the ATC environment, goals, tasks, and idioms. Aids to recall, such as video/audio recordings, should be used to the greatest extent possible to enhance the validity of the data. Selected research techniques and plans for site visits are designed to meet these conditions. Just as most of the research reviewed has used convergent methods (i.e., more than one particular technique), a combination of techniques to gain insight and enhance the validity of research findings will be used.

2.3 ANALYSIS AND SUMMARY OF METHODS.

Table 1 provides a summary of the methods and the measures evaluated against criteria for suitability. An "X" indicates that the particular method/measure satisfies a criterion. In some cases, placement of an "X" in a particular cell represents our particular orientation to these methods. For example, it might be argued that an unstructured interview will provide direct access to cognitive structures and strategies because the interviewer can ask about them. Our position is that findings based on unstructured interviews depend on the perspective of the interviewer and the operative definitions of cognitive structures and strategies. Different interviewers are likely to propose very different cognitive structures and strategies based on comments from controllers in unstructured interviews. Thus, unstructured interviews pose a reliability issue. Whether a particular technique provides direct access to cognitive structures and strategies also depends on the amount of additional analysis that is necessary as a basis for inference. If significant additional analysis is necessary, the technique does not provide direct access.

Suitable techniques for investigating strategies are retrospective verbalization, structured probed interviews, and video analysis. Pathfinder and MDS are among the appropriate statistical techniques for exploring the controllers' cognitive structures.

Table 1 Summary of Methods / Measures Evaluated Against Criteria

Technique	Non-disruptive of ATC Operations	Endorsed by Other Researchers	Preferred as Meaningful	Efficient Data Extraction	Direct Access to Knowledge Structures	Direct Access to Strategies
1. Concurrent Verbalization		X	X			X
2. Retrospective Verbalization	X	X			X	X
3. Unstructured Interview	X		(preliminary insight)			
4. Replay Technique	X	X	(in conjunction with interview)	X		X
5. Structured Interview with Probes	X	X	X		X	X
6. Video Analysis	X	X	X		X	X
7. Pathfinder	X	X	X	X	X	
8. MDS	X	X	X	X	X	
9. Structured Observation	X		(depends on observation tools)	(depends on tools)		X
10. Withholding Information		X	(provides ordering information)	X	X	X
11. Drawing Tasks	X	X	X	X	X	
12. Abstraction Hierarchy	X	X	X		X	
13. Logical Interference	X	X				
14. Critical Incident Technique	X	X	X	X		
15. Critical Decision Method	X	X	X	X		X
16. Cognitive Task Analysis	X	X	X			

Table 1. Summary of Methods / Measures Evaluated Against Criteria (Cont'd)

Technique	Non-disruptive of ATC Operations	Endorsed by Other Researchers	Preferred as Meaningful	Efficient Data Extraction	Direct Access to Knowledge Structures	Direct Access to Strategies
17. Dual-Coding Tasks	X	X	X	X		
18. Verbal-Cueing Tasks		X	X			
19. Paired Problem Solving	X					
20. Cognitive Style Assessment	X	X	X			
21. Simulator Problem Solving	X	X				
22. Literature-Based Analysis	X	X				
23. Memory Recall Tasks	X	X	X	X	X	
24. Cognitive Modeling	X	X				

3. CONCLUSIONS.

This report provides a broad overview of research techniques applicable to the study of air traffic controller knowledge and decision making. The methods reviewed have been evaluated for practicality and effectiveness for applied studies in air traffic control (ATC). A number of techniques, including retrospective verbalization, structured probed interviews, video analysis, Pathfinder, and multidimensional scaling (MDS), are among the appropriate methods for exploring controllers' cognitive structures.

REFERENCES

- Ammerman, H. L., Bergen, L. J., Davies, D. K., Hostetler, C. M. Inman, E. E., & Jones, G. W. (1988). FAA Air Traffic Control Operations Concepts (Vol. VI: ARTCC/Host En Route Controllers, Change 1, DOT/FAA/AP-87-01). Washington, D.C. U.S. Department of Transportation, Federal Aviation Administration.
- Bisseret, A. (1970). Memoire operationelle et structure du travail. Bulletin de Psychologie, 24, 280-294.
- Coeterier, J.F. (1971). Individual strategies in ATC freedom and choice. Ergonomics, 14(5), 579-584.
- Durso, F. T., Gronlund, S. D., Lewandowsky, S., & Gettys, C. F. (1991). Cognitive factors in the use of flight progress strips: Implication for automation (Cognitive Processes Laboratory Report). Norman, OK: Department of Psychology, University of Oklahoma.
- Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.
- Falzon, P. (1982). Display structures: Compatibility with the operators' mental representation and reasoning processes. In G. Johannsen (Ed.), Proceedings of the Second European Annual Conference of Human Decision Making and Manual Control (pp. 297-305).
- Federal Aviation Administration. (1987). Profile of operational errors in the national airspace system, calendar year 1986, Washington, D.C.: U.S. Department of Transportation, FAA.
- Federal Aviation Administration. (1990). The national plan for aviation human factors (Draft). Washington, D.C.: U.S. Department of Transportation/Federal Aviation Administration.
- Flanagan, J.C. (1954). The critical incident technique. Psychological Bulletin, 51(4), 327-328.
- Fraker, M. L. (1989). Attention allocation in situation awareness. Proceedings of the Human Factors Society (pp. 1396-1400).
- Johnson-Laird, P. N. (1981). Mental models in cognitive science. In D. Norman (Ed.), Mental models in cognitive science (pp. 147-191). New York: Ablex.

- Johnson-Laird, P. N. (1983). Mental models. Cambridge, MA: Harvard University Press.
- Klein, G. A., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. IEEE Transactions on Systems, Man, and Cybernetics, 19(3), 462-472.
- Klein, G. A. (1989). Recognition-primed decisions. In W. Rouse (Ed.), Advances in man-machine systems research (pp. 47-92). Greenwich, CT: JAI Press.
- Kruskal, J. B., and Wish, M. (1988). Multidimensional scaling. Beverly Hills, CA: Sage University Press.
- Lafon-Milon, M. T. (1981). Représentation mentale de la séparation verticale au cours du diagnostic dans le contrôle aérien: 3- Représentation des états futurs. (Rapport INRIA n°1. COR66). Le Chesnay, France: INRIA.
- Landis, D., Silver, C. A., Jones, J. M., & Messick, S. (1967). Level of proficiency and multidimensional viewpoints about problem similarity. Journal of Applied Psychology, 51, 216-222.
- Lapan, Y. A. (1985). Spatial representation and the activity of the air traffic controller. Vestnik Moskovskogo Universiteta Seriya 14 Psikhologiya, 14, 68-70.
- Leplat, J., & Hoc, J. (1981). Subsequent verbalization in the study of cognitive processes. Ergonomics, 24(10), 743-755.
- Means, B., Mumaw, R., Roth, C., Schlager, M., McWilliams, E., Gagne, E., Rice, V., Rosenthal, D., & Heon, S. (1988). ATC training analysis study: Design of the next-generation ATC training system. (OPM Work Order 342-036). Richmond, VA: Human Resources Research Organization (HumRRO International, Inc.).
- Mogford, R. (1990a). The air traffic controller's mental model and its implications for equipment design and trainee selection. Proceedings of the Canadian Conference on Electrical and computer engineering (Vol. II, 5.6.1.1-5.6.1.4). Montreal: Canadian Society for Electrical and Computer Engineering.
- Mogford, R. (1990b). The characteristics and importance of the air traffic controller's mental model. Unpublished doctoral dissertation, Carleton University, Ottawa, ON.

- Morrow, D., & Greenspan, S. (1989). Situation models and information accessibility. In N.E. Sharkey (Ed.), Models of cognition: A review of cognitive science (Vol. 1, pp. 53-77). Norwood, NJ: Ablex.
- Paivio, A. (1986). Mental representations: A dual coding approach. New York: Oxford University Press.
- Rasmussen, J. (1979). On the structure of knowledge - a morphology of mental models in a man-machine context. Roskilde, Denmark: Riso National Laboratory.
- Redding, R. E., Cannon, J. R., Lierman, B. C., Ryder, J. M., Seamster, T. L., & Purcell, J. A. (1990). Cognitive task analysis of prioritization in air traffic control. Report to the Federal Aviation Administration. McLean, VA: Human Technology, Inc.
- Redding, R. E., Ryder, J. M., Seamster, T. L., Purcell, J. A., & Cannon, J. R. (1991). Cognitive task analysis of en route air traffic control: model extension and validation. Report to the Federal Aviation Administration. McLean, VA: Human Technology, Inc.
- Roske-Hofstrand, R., & Paap, K. R. (1985). Cognitive network organization and cockpit automation. In R. S. Jensen (Ed.) Proceedings of the International Symposium on Aviation Psychology (pp. 71-78). Columbus, OH: Ohio State University.
- Roske-Hofstrand, R. J. (1989). Video in applied cognitive research for human-centered design. SIGCHI Bulletin, 21(2), 75-77.
- Rumelhart, D. E., & Norman, D. A. (1988). Representation in memory. In R. C. Atkinson, R. J. Hernstein, G. Lindsey, & R. D. Luce (Eds.), Stevens' handbook of experimental psychology (Vol. 2: Learning and cognition). New York: Wiley.
- Sanderson, P. (1990). Verbal protocol analysis in three experimental domains using SHAPA. Proceedings of the Human Factors Society 34th Annual Meeting (pp. 1280-1284).
- Schlager, M. S., Means, B., & Roth, C. (1990). Cognitive task analysis for the real(-time) world. Proceedings of the Human Factors Society 34th Annual Meeting (pp. 1309-1313).
- Schvaneveldt, R. W., Durso, F. T., & Dearholt, D. W. (1985). Pathfinder: Networks from proximity data. (Report No. MCCS-87-9). Las Cruces, NM: Computing Research Laboratory.

- Schvaneveldt, R. W., Durso, F. T., Goldsmith, T. E., Breen, T. J., & Cooke, N. M. (1985). Measuring the structure of expertise. International Journal of Man-Machine Studies 23, 699-728.
- Sperandio, J. C. (1971). Variation of operator's strategies and regulating effects on workload. Ergonomics, 14(5), 571-577.
- Sperandio, J. C. (1974). Extensions to the study of the operational memory of air traffic controllers. (RSRE Translation No. 518). Malvern, Worcs., England: Royal Signals & Radar Establishment.
- Vingelis, P. J., Schaeffer, E., Stringer, P., Gromelski, S., & Ahmed, B. (1990). Air traffic controller memory enhancement: Literature review and proposed memory aids (Report No. DOT/FAA/CT-TN90/38). Atlantic City International Airport, NJ: U.S. Department of Transportation, Federal Aviation Administration Technical Center.
- Whitfield, D. (1979). A preliminary study of the air traffic controller's picture. Journal of the Canadian Air Traffic Controller's Association, 11, 19-22, 25, 28.
- Whitfield, D., & Jackson, A. (1982). The air traffic controller's picture as an example of a mental model. In G. Johannsen and J. E. Rijnsdorp (Eds.), Analysis, design, and evaluation of man-machine systems: Proceedings of the IFAC/IFIP/IFORS/IEA Conference (pp.37-44). New York: Pergamon.
- Zachary, W., Ryder, J. M., & Zubritzky, M. C. (1989). A cognitive model of human-computer interaction in naval air ASW mission management (CHI Systems Technical Report 891215.8704). Spring House, PA: CHI Systems.

ACRONYMS

ARTCC	Air Route Traffic Control Center
ATC	Air Traffic Control
DYSIM	Dynamic Simulation
FAA	Federal Aviation Administration
FDEN	Flight Data Entry Notation
FPL	Full Performance Level
HTI	Human Technology, Inc.
HumRRO	Human Resources Research Organization
MDS	Multidimensional Scaling
Pfnet	Pathfinder Cognitive Network Analysis Tool
PVD	Plan View Display
SHAPA	Shell for Performing Verbal Protocol Analysis
SME	Subject Matter Expert