DOT/FAA/AM-01/16

Office of Aerospace Medicine
Washington, D.C. 20591

Planning in Air Traffic Control

Scott D. Gronlund
Michael R. P. Dougherty
Francis T. Durso
John M. Canning
University of Oklahoma
Norman, OK 73019

Scott H. Mills
SBC Technology Resources, Inc.
Austin, TX 78759

October 2001

Final Report

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

U.S. Department
of Transportation

Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-01/16		

4. Title and Subtitle Planning in Air Traffic Control	5. Report Date October 2001
	6. Performing Organization Code
7. Author(s) Gronlund, S.D.[1], Dougherty, M.R.P[1]., Durso, F.T.[1], Canning, J.M.[1], and Mills, S.H.[2]	8. Performing Organization Report No.
[1] University of Oklahoma Department of Psychology Norman, OK 73019 [2] SBC Technology Resources, Inc. Austin, TX 78759	10. Work Unit No. (TRAIS)
9. Performing Organization Name and Address FAA Civil Aerospace Medical Institute P.O. Box 25082, Oklahoma City, OK 73125	11. Contract or Grant No.
	13. Type of Report and Period Covered
12. Sponsoring Agency name and Address Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	14. Sponsoring Agency Code

15. Supplemental Notes

This research was supported by Contract #DTFA-97-G-037; Research Task #HRR-516.

16. Abstract

An experiment was conducted to examine the planning activities of en route air traffic controllers. Controllers were placed in the role of planners and verbalized a plan for controlling traffic to a tactician (another controller) who implemented it. Planning, which is typically tacit, was thereby made explicit by distributing it across these two individuals. Verbalizations from the planner to the tactician were coded and summarized. The direction of plan management and the degree of plan systematicity were influenced by the phase of the planning process and the predictability of the problem/environment. For the more-predictable problems, planning began with a bottom-up picture building phase followed by a top-down plan development phase. The less-predictable problems also began as a bottom-up process, but management of the subsequent phases was characterized by equivalent bottom-up and top-down contributions. In addition, planning was more systematic for the more-predictable problems. This understanding of the planning activities of controllers can serve as an antecedent to the development of computer tools to aid planning.

17. Key Words	18. Distribution Statement
Planning, Verbal Protocols, Air Traffic Control	Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	25	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGMENTS

We are grateful to our subject matter expert, Paul DeBenedittis (AMA-510), without whom this experiment would have been impossible. We are also grateful to Dick Pollock (AMA-510) for his cooperation. We appreciate the improvements to the manuscript that resulted from the comments of Paul Krois (AAR-100), Herbert Maier (Texas A&M University), Mike Mumford (University of Oklahoma), Julia Pounds (AAM-510), Dave Schroeder (AAM-500), and Scott Shappell (AAM-510). This research was supported by Federal Aviation Administration grant #DTFA-97-G-037 and was included in Research Task HRR-516.

Mike Dougherty is now at the University of Maryland, Department of Psychology, College Park, MD 20742; John Canning is now at Cambridge Research Associates, 1430 Spring Hill Road, Suite 200, McLean, VA 22102.

Please address correspondence to: Scott D. Gronlund; Department of Psychology; University of Oklahoma; Norman, OK 73019; e-mail: sgronlund@ou.edu.

Planning in Air Traffic Control

Where should we spend the night on our way to the Grand Canyon? For most families embarking on a summer driving vacation, they address questions like these before they set off on their trip—they make a plan. They gather information about when they can depart, how far they want to drive, when they would like to arrive, how much money they want to spend, and so on. They may try out various scenarios, or given various constraints, only one course of action may be available. But even the best laid plans often require modifications such as would be seen in our example should our family encounter bad weather, car repairs, or sleeping children. Regardless, planning is a complex and challenging activity that engages many and varied aspects of cognition (see Mumford, Shultz, & Van Doorn, in press, for a review), and encompasses many different domains (including vacation planning, see Stewart & Vogt, 1999).

Research on planning has examined three different environments in which we plan: simple and static, complex and static, and complex and dynamic. Byrne (1977) provided an example of planning in a simple, static domain. Byrne had six experienced cooks plan a three-course meal. The cooks talked aloud while they planned, yielding a verbal protocol. Verbal protocols are a common method of data collection in the study of planning, including the current study. Patterns in the protocols were classified to provide a description of the type of planning engaged in by the cooks. The planning was characterized as hierarchical: The cooks first set up a list of goals (e.g., main course, appetizer). Next, they subdivided the main goals into subgoals (main course should have a protein, a vegetable, etc.). Planning progressed in a hierarchically organized manner (i.e., the main course was selected before the dessert), until the meal was fully planned.

Planning differed in more complex, although still static, domains such as writing (Kellogg, 1994, Chapter 6) and errand running (Hayes-Roth & Hayes-Roth, 1979). Hayes-Roth and Hayes-Roth had participants plan a sequence of errands around town. Again, verbal protocols were collected, and they revealed a planning process that the researchers characterized as opportunistic. In contrast to planning meals, which proceeded in a top-down manner and involved successive refinements at lower levels, planning errands involved detailed sequences of specific actions being planned in conjunction with high-level modifications to the plan. Hayes-Roth and Hayes-Roth argued that planning in complex domains does not benefit from the discipline and structure imposed by a hierarchy because general solutions may not exist or may be computationally intractable. Although they were speaking of computer models of planning, computational intractability may also characterize the dilemma humans face in complex domains.

More germane to the present work are studies of planning in complex, dynamic domains. These include such domains as medicine (e.g., Kuipers, Moskowitz, & Kassirer, 1988; Xiao, Milgram, & Doyle, 1997), military operations (see Pew & Mavor, 1998), business (e.g., project management systems, Pietras & Coury, 1991), among others. Bainbridge (1997) characterized complex dynamic domains by the following four features: First, information may be unavailable, or ambiguous, regarding the state of some parts of the system. Consequently, operators need to search actively for information to keep ahead of the task, rather than merely react to it. Second, the operator must keep one or more independent entities under control, entities that will continue to change over time even if no action is taken. Third, these dynamic entities have several variables to be controlled that require that an operator allocate processing resources among several simultaneous tasks. Furthermore, because it is usually not possible to complete one subtask before starting another, subtasks must be interleaved (multi-tasking). Finally, it is impossible to anticipate all possible situations that might arise, which requires that operators adapt their behavior to the changing details of the situation or environment.

Air traffic control matches each of Bainbridge's (1997) four features. Controllers are responsible for maintaining separation among all aircraft in their sector (volume of airspace), and getting aircraft expeditiously to their destinations. Despite the radar and other equipment at their disposal, some information is unknown (e.g., exactly when a pilot begins to climb, how quickly the climb will be completed), while other information is ambiguous (e.g., will two converging aircraft be in conflict in 10 minutes?). Often times it is the complexity of the situation, coupled with time pressure, which results in information being unknown or ambiguous. The dynamic independent entities (the aircraft) continue to change position even if the controller issues no control actions, and at speeds of seven miles a minute, they do so rapidly. The aircraft are also characterized by more than one variable (minimally these include, aircraft speed, altitude, and heading). Transitioning between altitudes is especially problematic (Durso, et al. 1996; Gronlund, Ohrt, Dougherty, Perry & Manning, 1998). Finally, the situation is constantly changing as new aircraft enter the sector, the wind shifts, a runway closes, and so on.

Despite the widespread belief that operators in complex, dynamic domains engage in planning, research with air traffic controllers has usually characterized the activities of the air traffic controller as largely tactical or reactive in nature (Durso & Gronlund, 1999; Hutton, Olszewski, Thordsen, & Kaempf, 1997). Tactical decisions are those required for the resolution of immediate conflicts among a small number (2 to 3) of aircraft and are assumed to have a relatively short execution time (a few minutes). Klein (1989) proposed the recognition-primed decision (RPD) model to account for decision making of this sort.

Recent studies have demonstrated evidence of more strategic planning behavior if the controller was given sufficient time, and tools (e.g., Dougherty, Gronlund, Durso, Canning, & Mills, 1999; Moertl, et al. 2000). Strategic plans involve multiple aircraft over a relatively long period of time (up to 20 minutes). Future concepts have also proposed the establishment of a more strategic controller position (e.g., sometimes called a multi-sector D-side controller—N. Lawson & K. Thompson, personal communication, Dec., 15, 1997). Similarly, the NASA Ames Research Center has proposed creating an "airspace coordinator" position (Vivona, Ballin, Green, Bach, & McNally, 1996). These positions have one person responsible for a multiple-sector airspace, making planning decisions about traffic in the sectors, and delegating responsibility for tactical decisions to sector-level controllers (who likely will maintain final authority over the proposed solution). More needs to be learned about the planning behavior of air traffic controllers so that recommendations can be made regarding how best to implement more strategic air traffic control, and what tools can best support that planning (see Canning, Johansson, Gronlund, Dougherty & Mills, 1999).

The present study examines two aspects of opportunistic planning: the direction of plan management and the degree of plan systematicity. Behavior would be guided from the top-down if it were coordinated by a plan, guided from the bottom-up if it were driven by events in the environment, or bi-directional if it involved equivalent contributions from both. A plan would be considered systematic if the opportunities could be anticipated but unsystematic if transitions between behaviors tended to be reactive. What does a plan look like in the air traffic control domain?

The direction of plan management and the degree of plan systematicity varies in different complex domains. Hayes-Roth and Hayes-Roth (1979) found that, in the complex but static domain of errand planning, high-level and low-level aspects of planning (bi-directional management) competed in an unsystematic manner. In other words, you might operate at the detailed level when looking for the errand closest to your current location, but subsequently make a decision at a more abstract level when you discover a cluster of errands that can be conducted nearby. Later, you might discover that one of the errands in the cluster can also be conducted near your final destination, which results in a move back to a detailed level of abstraction, and a decision to plan backward in time from the final destination (example taken from Hayes-Roth & Hayes-Roth, p. 282). In contrast, Johannsen and Rouse (1983) argued that in the complex but dynamic domain of a piloting an aircraft, planning behavior was often directional and systematic—organized from the top-down, with interrupts caused by unanticipated, but principled, events. For example, a pilot following a script for descent might be interrupted by the report that the runway was closed. However, the descent script remains active while the pilot seeks out an alternative. If an alternative airport is nearby, that can provide the opportunity that allows the plan to proceed.

On the surface, it would seem that plans in the air traffic control domain should resemble those that describe the plans of a pilot. However, as Johannsen and Rouse (1983) themselves pointed out, characteristics of a plan depends on the events that occur and whether or not they can be anticipated, and the phase of the planning process.

The ability to anticipate events depends on the predictability of the environment, which should affect plan systematicity. A less-predictable environment results in more information being unavailable or ambiguous, perhaps because it is too costly in terms of time or mental effort to anticipate problem states. Difficulty in anticipating situations may require an operator to adapt their behavior to the situation more frequently, resulting in less systematic planning. On the other hand, problem states are easier to anticipate in a more predictable environment. The greater degree of predictability can allow an operator to anticipate what information is important, and when it will be important, which may allow for more systematic planning.

Phase of the planning process should affect the direction of plan management. That is, an operator might allocate processing resources differentially during different phases. A bottom-up component may be central during the initial stages of plan development as information from the environment is collected. Bottom-up planning explicitly acknowledges the role of the environment and the contribution of situation assessment to the development of a plan. Durso and Gronlund (1999) argued for the importance of the environment in managing a complex dynamic system (see also Adams, Tenney, & Pew, 1995). The importance of the environment was also embodied in Brunswikian theories of decision-making (e.g., Gigerenzer, Todd, and the ABC group, 1999; Hammond, Hamm, Grassia, & Pearson, 1997; Kirlik & Bisantz, 1999).

Planning would be characterized as being top-down if a pre-stored plan were used as the organizing structure for comprehending the situation and controlling the traffic. For example, a planner working in a familiar sector at a routine time of day might be able to anticipate the traffic flow. She might know that certain flights enter her airspace at specific times, and therefore can make provisions for these aircraft in anticipation of their arrival. Such a plan would be predetermined insofar as the default values of the plan can be specified by the controller's knowledge of the airspace and the typical traffic patterns.

Planning would be bi-directional (exhibit a balance between a top-down and a bottom-up contribution) if no single overarching plan could be developed to coordinate the entire situation. Consequently, the situation would be parsed into small, more manageable, units (many mini-plans rather than one overarching plan); the direction of management shifting between top-down when a mini-plan is being put in place, to bottom-up when data collection is required prior to developing the next mini-plan.

Overview of experiment

The focus of the present experiment was on en route air traffic controllers. These controllers are responsible for the high-altitude, high-speed portion of a flight. En route controllers are responsible for a volume of airspace called a sector. Their job is to route aircraft safely and expeditiously through the sector to the next sector. A team of two controllers often staffs each sector, especially when the airspace is busy. The controller sitting in front of the radar display and talking to the pilots is called the Radar-side, or R-side controller. The primary tool is a 2-D radar display with a radar target for each aircraft. Each radar target has an attached data block of information that includes the call sign, ground speed, and altitude. The Data-side, or D-side controller, sits in front of the flight progress strips. There exists one (or sometimes more) flight progress strip for each aircraft on the radar display. Flight strips, typically, are 20 x 3-cm rectangular pieces of paper that include a variety of information regarding specific aircraft, including the call sign, aircraft type, requested altitude, requested speed, route of flight, etc. The D-side controller assists the R-side controller by updating the strips (e.g., as planned altitude becomes actual altitude). The D-side controller may also provide a second set of eyes on the radar, and may be called upon by the R-side controller to do some preplanning. However, final authority over any plans, their execution, and responsibility for their implications, falls on the R-side controller.

We examined planning by air traffic controllers in less- and more- predictable (or structured) air traffic control problems. The less-structured problem included aircraft en route to many different destination airports. The job of the controller in this situation was to move these aircraft through the sector safely and expeditiously. Most aircraft were *crossing* one another's route of flight. The more-structured problem included aircraft coming from many different locations but needing to be *sequenced* to land at the same destination airport. To enhance further the difference between the two types of problems, we also varied the routing of the aircraft. Currently, most aircraft fly along routes (so-called "highways in the sky”). This was the case in the sequencing problems. As a result of flying along established routes, there are particular points in a controller’s airspace where routes intersect and merging aircraft will conflict (like intersections at street corners). This was not the case in the crossing problem, which simulated some aspects of free flight [1] (FAA, 1995; RTCA, 1995). Most aircraft in the crossing problem flew direct or straight-line routes through the sector (something that has already been implemented by the FAA at high altitudes). This meant that aircraft could intersect at any point in the airspace. As Carlson, Rhodes, and Cullen (1996) argued, direct routing is likely to result in a significant increase in the amount of tactical separation actions and a corresponding decrease in strategic (longer-term) separation actions.

The sequencing problems granted more predictability than the crossing problems. The variety (and unfamiliarity) of intersection points in the crossing problem would make it more difficult for the controllers to anticipate what was going to happen. As a consequence, we hypothesized that sequencing problems should exhibit more systematic planning, and conversely, the crossing problems should exhibit less systematic planning. Planning was also tracked through three distinct phases of the experiment, ranging from the first look at the problem, through the development of an initial plan, to the maintenance of that plan as the situation evolved. The planning process should progress from bottom-up management as the planner is building their picture of the situation, to top-down management as a plan is put in place and maintained.

METHOD

Participants

Twelve en route air traffic controllers participated in this research study. All were instructors at the FAA Academy and were familiar with the AeroCenter airspace used in the experiment. All were full-performance level controllers, which meant that they were certified to work a sector independently, in contrast to a trainee who must work with a full-performance level controller. They had been full-performance controllers for an average of 6.3 years. They last worked in the field an average of 2.5 years, with a range of 0.5 to 6.2 years.

Materials

The experiment was conducted at the Radar Training Facility at the Mike Monroney Aeronautical Center in Oklahoma City. The Radar Training Facility has high-fidelity air traffic training simulators used to provide radar training. Communications between the controllers and the aircraft took place in the same manner as in the field, although the aircraft were “piloted” by ghost pilots who controlled the simulated aircraft based on the controller’s instructions.

The equipment consisted of the circular radar display (the Plan View Display), two bays of paper strips, the Computer Readout Display, and a keyboard and trackball. The Plan View Display indicates the 2-D position of the aircraft with an attached data block containing information including the aircraft’s call sign, altitude, and ground speed. One flight progress strip for each aircraft was stacked vertically in a strip bay adjacent to the radar display. The planner could call up an aircraft’s flight plan on the Computer Readout Display.

Three scenarios were developed with the help of our Subject Matter Expert. The scenarios were judged by the Subject Matter Expert to exceed the workload level typically experienced in the field by a team of two controllers. Two of the scenarios were designed to be very similar. In both, the primary problem to be solved involved *sequencing* aircraft for ultimate landing at Dallas/Fort Worth airport. For this reason, the strip bay was organized with a column of Dallas/Fort Worth arrivals and a column of “other” aircraft. Within the two columns, the strips were organized by increasing time relative to Tulsa (i.e., all strips included information on the expected time that they would pass over Tulsa). This set-up facilitated planning and was the first thing subjects did in the practice sessions. In the third scenario, aircraft were *crossing* traffic for one another (e.g., an aircraft heading north and an aircraft heading west might cross over Tulsa). In addition, because aircraft flew direct, straight-line routes through the sector, they could

[1] Free flight has other implications for the air traffic controller, including decentralized decision-making power and changing the controller’s role to be more of a monitor, that were not examined in the present experiment. The advantage of free flight, from the perspective of the airlines and their passengers, is shorter, more direct routes between departure and destination airports.

intersect at any point in the airspace. Although we believe that it was the crossing traffic more so than the use of direct routing that made this scenario less predictable, additional research should be conducted to determine the contribution of each factor.

There were six counterbalanced orders of scenarios. This resulted from randomly selecting one of the two sequencing scenarios to be third and then randomly assigning the other sequencing scenario to be either first or second, with the crossing scenario filling the remaining position. The 12 participants allowed us to rotate through the six counterbalanced orders twice. One of the sequencing scenarios was always last. This was because we closed the Dallas/Fort Worth airport midway through the last scenario. Closing the airport provided a partial replication of the experiment because it necessitated the development of a new plan. No prior warning was given about closing the airport.

Procedure

Tactics and planning are normally confounded because both types of decision making often lie within the same head, even when a team has responsibility for a sector. Roughly speaking, planning occurs further in advance than tactics and involves a larger number of aircraft. Tactics occur in the current moment and typically involve the separation of pairs of aircraft. Although the D-side controller does engage in preplanning (along with several other activities), the R-side controller, the same person who verbalizes tactical actions, also decides upon the actual plan that will be conveyed to the pilots. This natural confounding of planning and tactics led to the development of the distributed verbal protocol method. The distributed verbal protocol method made explicit that which is tacit, by distributing the cognition across individuals. In other words, we put the role of the planner and the tactician (the implementer of the plan), into different heads.

Although some aspects of verbal protocols and their analysis are controversial (e.g., Wilson, 1994), Payne (1994) argued that "best practices" are developing. These include emphasizing the primary task over the protocol task, collecting on-line rather than retrospective protocols, requiring that operators verbalize thoughts rather than analyze those thoughts, and minimizing the social component. The particular methodology we employed adopted many of these "best practices." By requiring the planner to verbalize his or her plan to the tactician (another expert controller), we believe we were able to capture a relatively pure description of controller planning. We base this assumption on the fact that the primary task and the verbalization task were the same thing for our planners; their job was to articulate the plan for the tactician. The protocol collection was concurrent with their planning and did not require that the controller analyze their verbalizations. Furthermore, the planner described the plan to the tactician (another expert who spoke the same technical language), not the experimenters (non-experts). The same expert controller served as the tactician (the Subject Matter Expert) for all 12 experimental participants. Finally, it is worth noting that the planners found the "division of labor" required by the task to be straightforward and easy to implement.

The tactician's actions were directed at solving problems that might arise in the relatively near future (0 – 5 minutes) and which had relatively little long-term impact. For example, two aircraft on conflicting trajectories must be separated before they violate each other's airspace. If the resolution of a conflict had no direct impact on future control (i.e., no matter what control action was taken to resolve the conflict, a new problem would not be created) the action was said to be tactical in nature. In contrast, the planner's efforts were to be directed at solving problems that might arise in the relatively distant future (5+ minutes) and which had a relatively long-term effect. For example, six aircraft landing at the same airport need to be sequenced to minimize delays to any single aircraft, and to maintain separation between aircraft. These types of decisions required that the controller have a broader plan that incorporated all six landing aircraft, as well as any crossing aircraft that might disrupt the sequence.

The instructions began by delineating the roles of the tactician (the Subject Matter Expert) and the planner (the participant). The tactician's job was to maintain separation between aircraft; he would make whatever altitude, speed, and heading changes necessary to maintain separation. The planner's job was to make the tactician's job as easy as possible, and to give the tactician a plan for managing the flow of traffic in the sector. The planner was told to utter tactical control actions only if it was necessary to implement their plan. Because the planner was not allowed to hear pilot requests, the tactician relayed them and all requests had to be granted by the planner.

Prior to beginning, a minor addition to the AeroCenter airspace was pointed out. The planner was also given a list that included all the airlines used in the scenarios, along with their call letters and designators. A lengthy practice session followed. A scenario was completed with several key events embedded in it that served to illustrate, in a concrete way, the respective roles of the planner and the tactician. For example, a simple, isolated, two-aircraft conflict was offered as an example of a tactical conflict. No matter what solution to the conflict was implemented by the tactician, these aircraft would not affect the overall plan. The planner was told that this was the kind of conflict that could be pointed out if they wished, but it could be left to the tactician to solve because it did not matter how it was solved. Another example involved a two-aircraft conflict that required a particular solution (i.e., some tactical solutions to the problems affected the plan). Similarly, a third example involved another two-aircraft conflict, but some tactical solutions created a conflict with a third aircraft. In this situation, the planner was told that we were looking for them to provide a single solution that would solve both problems simultaneously. After the practice scenario was completed, the participants read a brief summary of the instructions.

ðThe planner and the tactician each wore a microphone. The tactician sat in front of the Plan View Display (PVD). The planner usually sat in front of the strip bay, although he or she could move around behind the tactician if desired. The planner was given full access to all the functionality of the PVD and Computer Readout Display (CRD). In particular, this included calling up an aircraft's destination using the CRD, displaying the route of flight of an aircraft on the PVD, or projecting a straight-line vector for all aircraft 1, 2, 4, or 8 minutes into the future.

The planner was given a sheet with the altimeters, the traffic flow (either the Dallas rush was coming in the sequencing problems, or the aircraft were on direct routings in the crossing problem), and flow restrictions (10 miles between aircraft heading to Dallas/Fort Worth, 5 miles is normal). The planner then sat down and was asked to describe what he or she saw and to verbalize the plan. The tactician took notes on the plan because he would have to implement it. The scenario began once the planner was finished.

The experiment began with the scenario paused after it was discovered during the practice runs that if the scenarios were active, the planner immediately fell into a tactical mode until he or she could determine what to do. In the field, a controller who was taking over for another controller would receive a position-relief briefing that would provide for them the "big picture" of what was going on before the relieving controller took over responsibility for the sector. Because an examination of the position-relief briefing was not of interest in this experiment, we decided to pause the scenario and give the planner as much time as desired to develop the plan. Once the planner was ready, the scenario began. The planner was instructed to modify the plan as necessary. At the 10-minute mark, the scenario was again paused and the planner was again asked to describe what he or she saw and to verbalize the plan. The tactician also completed a short questionnaire. The scenario then ran for another 10 minutes, after which the tactician and planner each completed a questionnaire.

This was the procedure for the first two scenarios, with a 15-minute break following each. The third scenario, which was always a sequencing scenario, began like the others, with the following change at the 10-minute mark. After pausing the scenario, we handed the planner a note that stated that an accident had occurred at Dallas/Fort Worth and one half of the Dallas/Fort Worth arrivals had to be rerouted to Oklahoma City (via a certain airway, at or above a certain altitude, with 10 miles separation between aircraft) and one half had to be rerouted to Houston (similar restrictions). The planner was asked to describe the new situation and verbalized a new plan for dealing with it. After completion of this third scenario, a final questionnaire collected biographical information. The participants were then debriefed and released. The entire experiment took approximately 2.75 hours per participant.

RESULTS AND DISCUSSION

The primary data collected were the verbalizations from the planner to the tactician. (All the planner-tactician teams safely and expeditiously moved the traffic through the sector with no separation violations. Dougherty et al. [1999] summarized other aspects of the data.) The first step in analyzing these data involved subdividing the verbalizations into idea units. A coding scheme was then developed that classified each idea unit into one of four major categories.

Coding scheme

The major and minor categories used in this study are presented in Table 1 and briefly summarized here. The first major category, Collect Data, was used whenever the planner identified a piece of information without performing an action on it. The two subcategories were DATa (information came directly from a source) and INFerence (information inferred from the problem or from domain-specific knowledge). An example of DAT would be "AAL123 goes to Dallas/Fort Worth," which could be read from a strip; an example of INF would be "DAL123 is an overflight," because that information was deduced by determining that an aircraft neither departed from nor landed in your sector.

The second major category, Monitor (M), was used whenever the planner re-checked what he or she had already done, or recognized that specific aspects of the situation were *not* a problem. M had four subcategories: ENVironment—when the planner examined jetways, airports, or altitudes (e.g., "Flight level 330 [33,000 feet] is available"), PLan, when the status of the plan was monitored (e.g., "looks like DAL123 and AAL123 have sufficient spacing"), TACtics, when the planner monitored the control actions of the tactician (e.g., "AAL123 will need more of a vector" in response to the tactician having issued a heading to AAL123), AirCraft, when the planner checked on the status of an aircraft (e.g., "DAL123 is looking good").

Whenever a planner identified a problem, the third major category, Problem(P), was used. It had two subcategories: The controller could IDentify a specific problem or DETect a problem without identifying its exact nature. An example of the former would be "AAL123 is overtaking DAL123"; an example of the latter would be "DAL123 and AAL123 is not going to work."

Finally, there were times when the planner was constructing plans for actions to take in the future. In these instances, the Develop Plan (DP) category was used. Develop Plan had two subcategories: A future action could be either an a:Unconditional Action (UA) or a Conditional Action (CA). The former should be carried out regardless (e.g., "take DAL123 direct Amarillo"), the latter had some temporal, spatial, or logical constraint placed on it (e.g., "after DAL123 passes Tulsa, take him direct Amarillo").

Using this coding scheme, the first two authors independently coded each idea unit extracted from the planner verbalizations. To check on the level of agreement, we compiled the data from three randomly chosen subjects for each of the three scenarios. Agreement was high, with no notable differences as a function of type of problem, phase, or category being coded (overall agreement 81%, kappa .75, $z = 14.15$, $p < .001$, signaling that the level of agreement was significantly greater than expected by chance). Disagreements were resolved in the following manner: If

Table 1

List of categories and their definitions

Major category	Subcategory	Definition
Collect Data (CD)	Data (DAT)	Information read-off a source
	Inference (INF)	Information inferred from the problem or from domain-specific knowledge
Monitor (M)	Environment (ENV)	Aspects of the environment examined and no problem found
	Plan (PL)	Plan examined and no problem found
	Tactics (TAC)	Tactics examined and no problem found
	Aircraft (AC)	Aircraft status examined and no problem found
Problem (P)	Identify (ID)	Problem identified
	Detect (DET)	Problem detected but not identified
Develop Plan (DP)	Unconditional action (UA)	Future action issued, no constraints
	Conditional action (CA)	Future action issued, constraints

one of the raters evaluated an entry as NONE and the other gave it a code, we used the code. In cases where both raters assigned different codes to an entry, the item was discussed and the difference was resolved.

The different phases of the experiment help organize the analyses to follow, beginning with the verbalizations that took place when the scenarios were paused. For reasons that will become apparent, this initial planning period was split into two parts—a picture building phase and a plan development phase. Examination of these two phases was followed by an examination of the verbalizations that took place during the first 10-minute period in which the scenario was running. This is referred to as the plan maintenance phase. Recall that prior to the second 10 minutes of the second sequencing scenario, which was always the last scenario in the experiment, the scenario was paused and the planner was informed that the Dallas/Fort Worth airport was closed. The period during which this second sequencing scenario was paused was referred to as the replan phase. The subsequent 10 minutes of running this "new" problem was called the replan maintenance phase.

The data were analyzed at three levels of analysis. The first focused on the four major categories and the proportion of time that they occurred in each phase of the experiment. It was because of these data that the initial planning period was split into a picture building phase and a plan development phase. The next level examined the latent structure within the transition matrices of the coded planner verbalizations for all 10 subcategories using the Pathfinder scaling algorithm (Schvaneveldt, 1990; Schvaneveldt, Durso, & Dearholt, 1989). The resulting graphs can be used to generate the verbalization strings typically produced by planners in the various conditions. These strings represent the types of transitions between verbalizations that were, and were not made and as such, they depict a planning grammar. The final level of analysis involved various summary measures extracted from the Pathfinder graphs. The summary measures informed us about the direction of plan management and the degree of plan systematicity.

Picture building and plan development (scenario paused)

The first compilation of the data examined rolling blocks of 10 verbalizations for the major categories only. In Figure 1, the first point on the x-axis gives the proportion of verbalizations in the four major categories among the first 10 verbalizations; the second point on the x-axis gives the same for verbalizations 2 through 11, and so on. This method is a generalization of data smoothing, although a span of 10 verbalizations is larger than what is typically used (see Velleman & Hoaglin, 1981). It was necessary because averages of four variables were being computed simultaneously. The use of rolling blocks was preferable to dividing the data into discrete intervals because it better conveyed the continuous nature of the planning task. Furthermore, dividing the data into discrete intervals would change none of the conclusions (as the reader can judge by looking at intervals 0, 11, and 22, or 1, 12, and 23, which have no overlapping data).

The top panel gives the data for the crossing scenario and the bottom panel for the two sequencing scenarios (averaged together). In both the crossing and sequencing scenarios, the proportion of collect data (CD) initially exceeded the proportion of develop plan (DP), but at some point this relationship reversed. It was because of this reversal that two distinct subphases were hypothesized to occur during this phase. Because each individual's graph crossed at a different point, participants were lined up according to the point that CD first exceeded DP, and then tallied backward and forward from the different crossing points as far as the data allowed (i.e., until participants ceased contributing data).

In both scenarios, planners began by collecting data about the situation to "build the picture." They had to build their picture before they could begin to develop a plan to control traffic.[2] In the sequencing scenario, a trade-off between data collection (CD) and plan development (DP) was the primary activity. In the crossing scenario, the picture building phase was less distinct as CD traded-off with DP, monitoring (M), and problem (P).

[2] It is possible that we found more evidence for a reliance on data collection than would exist in the field: The planners were familiar with the airspace, but not the aircraft involved. In contrast, in the field a controller might know that AAL123 was arriving from Kansas City and DAL456 from Memphis at 11:00 and both were heading to Dallas/Fort Worth, because it happened every day.

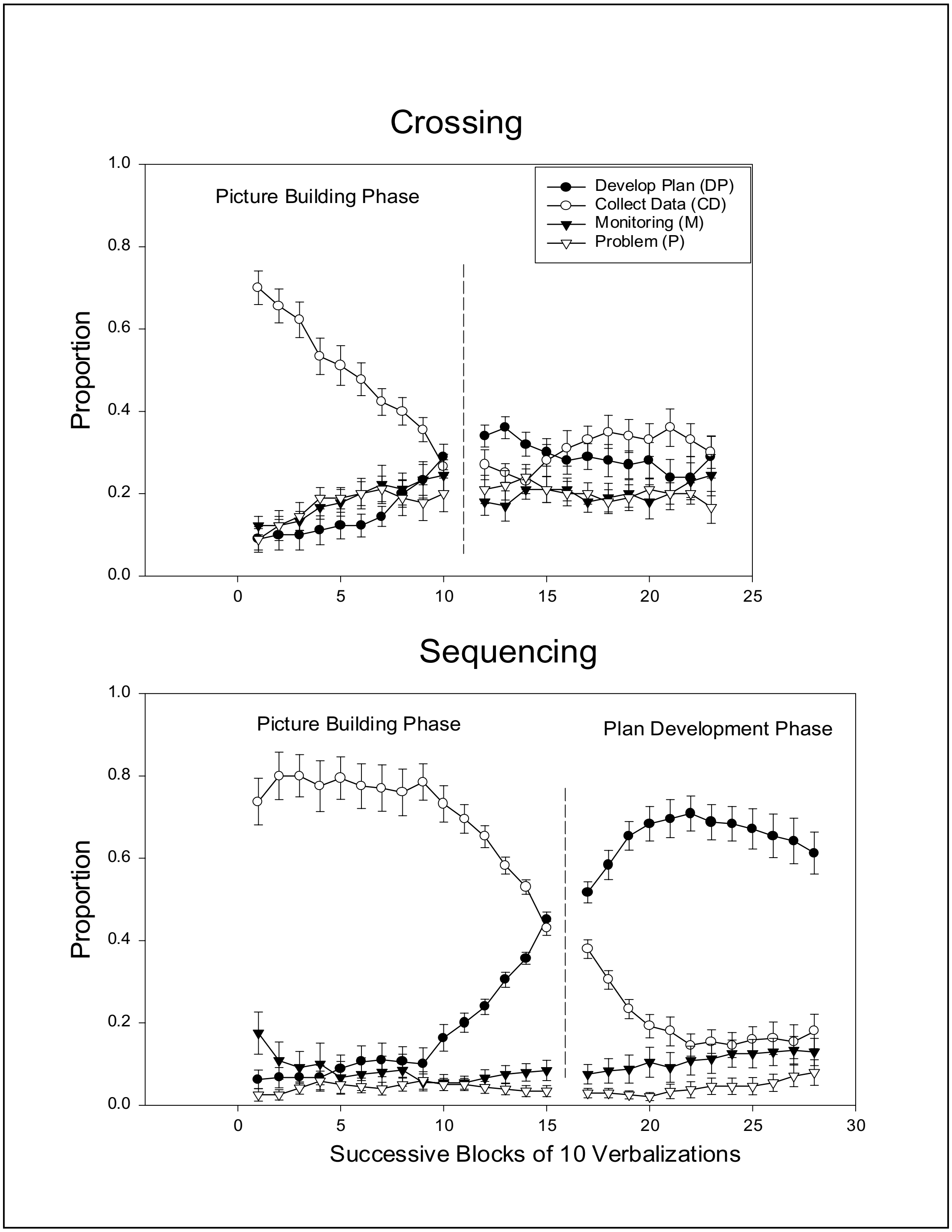

Figure 1. Proportion of verbalizations for the four major categories as a function of rolling blocks of ten verbalizations. Crossing scenario on the top and Sequencing scenario on the bottom. The gap in the middle of the graph shows the point where DP≥CD. Error bars show one standard error of the mean. The scenario was paused during this phase.

After the point that DP first exceeded CD, there was a striking difference between the sequencing and crossing scenarios. A plan development phase was evident in the sequencing scenario; as data collection decreased, plan development increased. Monitoring and problem verbalizations were rare. The situation was markedly different in the crossing scenario. There was no distinct plan development phase, perhaps because no single plan existed to coordinate activities. Plan development was no more likely than data collection during this phase. Monitoring (M) and problem verbalizations (P) were also more frequent in the crossing scenario than they were in the sequencing scenarios.

CD and DP were the most frequent verbalizations overall, and were clearly contingent on time. However, once the initial picture building phase was completed, the more-predictable sequencing scenarios afforded greater opportunities for integrating aircraft into a plan. In the crossing scenario, on the other hand, there were fewer DP verbalizations and more M and P verbalizations, perhaps because the crossing problems restricted how much planning could be done. It was more difficult to anticipate problem states in less-structured problems, and planners had to spend more time reacting to the ever-shifting situation.

The above analyses examined the proportion of each type of major category verbalization. However, it is also useful to examine, at a more detailed level, how participants transitioned from one type of activity to another. The structure of these transitions was informative regarding the degree of plan systematicity. It also indicated whether planning was guided in a top-down manner (developing the plan initiated the search for information from the environment), from the bottom-up (searching for information triggered plan development), or was a balance between these two.

The strings of coded verbalizations were summarized as transition matrices between one verbalization and the next. For this and all subsequent analyses, we considered the subcategory level, which resulted in a 10 x 10 matrix of transitions. The transition matrices were normalized by the total number of verbalizations uttered by that participant before averaging across participants. Four matrices were produced—verbalizations occurring before or after DP first exceeded CD (picture building and plan development), crossed with the two types of scenarios (crossing and the two sequencing scenarios averaged together).

Vortac, Edwards, and Manning (1994) suggested weighting the transition between successive behaviors as an exponential function of the temporal interval between the behaviors. That would mean that behaviors that occurred in close temporal proximity would "count more" than would two behaviors that did not occur in close temporal proximity. We did not exponentially weight the transitions because the planners talked almost continuously, at least when the scenario was paused, and the majority of the data came from quite compressed time spans. Transition matrices consisted of proportions, rather than raw frequencies, to ensure that each participant had the same impact, on the average, irrespective of how verbose they were. However, participants might have talked more because they had a more liberal criterion for what they would verbalize. For example, one participant might verbalize A → B → C while another might only verbalize A → C. Averaging across participants has the advantage of eliminating idiosyncratic patterns of verbalizations in favor of those produced by a preponderance of the participants. This was appropriate, given that we were trying to describe typical planning performance.

Many procedures have been developed to reveal the latent structure underlying a set of data like these transition matrices (e.g., multidimensional scaling—Shepard, 1962; clustering—Johnson, 1967). These procedures share the assumption that the observed data reflect latent ("true") structure plus statistical noise, and that the two can be separated by mathematical means. We chose Pathfinder because, in contrast to multidimensional scaling, Pathfinder can represent asymmetric transitions, which are likely in the present experiment. Pathfinder reduces a transition matrix to a graph by eliminating those transitions between verbalization categories that do not satisfy metric properties (e.g., the triangle inequality). The k transitions chosen will be the shortest distance between all events i and j, given k transitions, which means that every link in the resulting graph is a link on some minimal path between two nodes. A family of graphs can be created, depending on the metric used to compute path distance. We chose parameter values that created the sparsest Pathfinder graphs (i.e., the minimum number of links): q was set to 9 (10 verbalization categories minus 1), and r (the value of the Minkowski distance metric) was set to ¥.

The graphs summarized the structure of the verbalizations from the various conditions. The two sequencing scenarios (averaged together) are across the top of Figure 2; the crossing scenario is across the bottom. The picture building phases are on the left, the plan development phases are on the right. The structural similarity between a pair of graphs was assessed by comparing the number of links two graphs shared normalized against the total number of links (see Goldsmith & Davenport, 1990 for discussion of various ways to compare graph similarity). The picture building and plan development sequencing graphs (the top two in Figure 2) share only five links. This provided additional evidence for the distinct planning phases identified in the sequencing scenarios; all of the other comparisons of structural similarity between pairs of graphs shared at least nine links.

The size of a node was directly proportional to the proportion of all verbalizations of that type within a given phase. For example, .72 of all verbalizations in the sequencing scenarios in the picture building phase were CD-DAT; therefore, the CD-DAT node represents 72% of the total area of all the nodes depicted in the top left-hand graph of Figure 2. An arc connected verbalizations that tended to co-occur. The thicker a link, the more frequent the transition. Note that some of the links were loops—a verbalization of a particular type followed by another one of the same type. To enhance readability, we made an effort to keep the nodes in the same location across graphs (Monitoring on the left side, Develop Plan near the top center, Problems near the bottom, and Collect Data near the center). As a result, some links must pass through a node. To enhance readability, these links were drawn with dashed lines. The distance between nodes in the network is irrelevant.

The resulting graphs function as a planning grammar because they can be used to generate the types of transitions between verbalizations that were, and were not, made. For example, a planner in the picture building phase of the sequencing scenarios usually began with a verbalization involving some data collection (CD-DAT). This would typically be followed by several more CD-DAT's (the loop from CD-DAT to CD-DAT). The planner had to first determine where the aircraft were going to determine which aircraft needed to be sequenced to Dallas/Fort Worth, and which ones did not. At some point, a piece of the plan would be constructed (DP-UA), after which the planner returned to more data collection (CD-DAT). A DP-UA verbalization was seldom if ever followed by anything besides another CD-DAT (hence the link back to CD-DAT and no links to anything else). Interspersed sporadically among these verbalizations were other verbalizations involving monitoring or a problem. These verbalizations were rare; they were usually triggered by a preceding CD-DAT verbalization and, typically, were followed by a CD-DAT verbalization.

These graphs are a rich source of hypotheses about the planning process. Many implications can be drawn from them that can be subjected to future testing. A couple of implications will be illustrated. One difference between the picture building phases of the two scenarios involved monitoring. For instance, the identification of a problem (P-ID) preceded monitoring the plan (M-PL) in the crossing scenario, but M-PL was triggered by data collection (CD-DAT) in the sequencing scenario. This might signal that problems in the crossing scenario involved the plan, but acquiring the data was the problem in the sequencing scenario. In the plan development phase for the sequencing scenario, monitoring the plan (M-PL) was part of the plan development process because it was preceded and followed by DP-UA. However, in the crossing scenario, data collection (CD-DAT) triggered M-PL. This is consistent with a more-predictable sequencing scenario; the planner could rely on the internal environment (their mental simulation of the plan) rather than needing to refer to the external environment for verification regarding the implications of their plan.

The following guidelines were used to interpret the graphs. The focal node of a Pathfinder graph signaled the direction of management. A graph whose focal node was data collection (i.e., CD-DAT) was characterized as bottom-up. A graph whose focal node was developing the plan (i.e., DP-UA) was characterized as top-down. A graph that contained bottom-up and top-down foci was characterized as bi-directional. The focal node of a Pathfinder graph was determined by computing its median (see Durso, Rea, & Dayton, 1994). The median of a graph was the node that had the smallest average distance to all other nodes. For example, a hospital should be placed at the median of a city because you want to minimize the average travel time for all residents, not just those that live in the largest apartment complex. In other words, medians are based on linkages, not on the frequency (size) of the nodes.

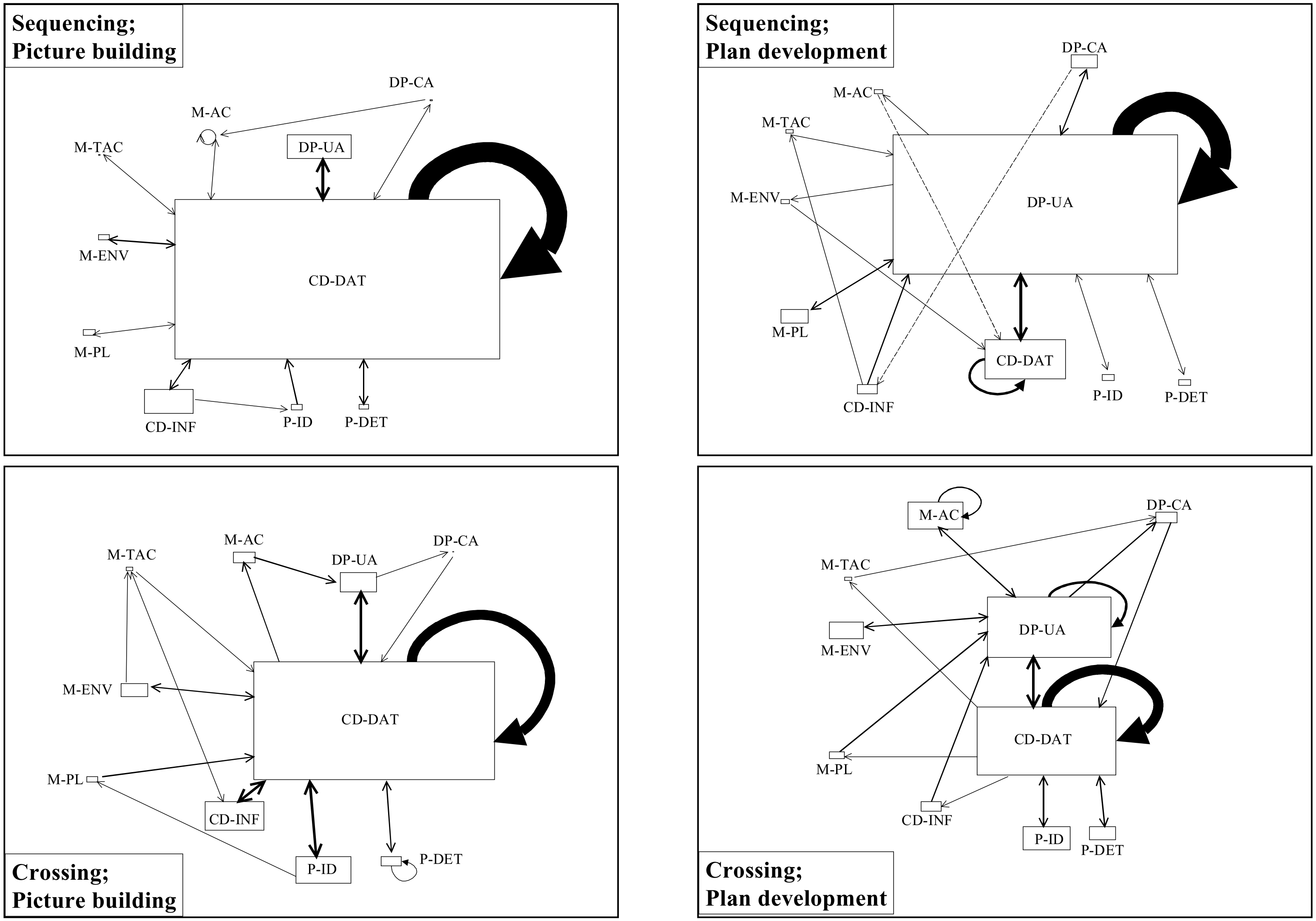

Figure 2. Pathfinder graphs for $q = n - 1$ and $r = \infty$. The Sequencing scenarios (averaged together) are across the top, the Crossing scenario is across the bottom. The Picture Development phase is on the left, the Plan Development phase is on the right.

The median can represent either the most influential or the most prestigious node. The influence median was the node from which every other node could be reached in the minimal number of transitions; the prestige median was the node that every other node could reach in the minimal number of transitions. In these data, the influence and prestige medians were identical. Consequently, Table 2 gives only the influence medians for the two scenario types crossed with the various phases of planning. The actual median values (the number of transitions) are given in parentheses. For example, an influence value of 1.1 indicated that 1.1 transitions were needed to get from the median node to any other node. The median of the picture building phases of both the sequencing and crossing graphs was CD-DAT, an indication that the initial phase of planning was best described as a bottom-up process. In the plan development phase of the sequencing problem, DP-UA usurped the role CD-DAT played in the picture building phase, an indication that this phase of planning was managed from the top-down. The situation was different in the crossing scenario: There were two medians, DP-UA and CD-DAT, signaling bi-directional management of planning.

In a systematic plan, a particular node should orchestrate the flow of activity, and disruptions to the flow should involve only a single-step departure. A less systematic plan, on the other hand, would be one in which the flow of activity changed course frequently, reacting to the ever-changing situation. Switching between tasks should occur more often and involve more intervening activities. For example, data collection might trigger monitoring of an aircraft, which might trigger problem identification, which might lead to plan development, which might trigger another problem identification, and so on. Plan systematicity was assessed using the number of unique cycles in a graph. A cycle is a sequence of three or more verbalizations (excluding loops) that did not repeat the same verbalization. For example, DP-UA to M-AC to DP-CA was a cycle; DP-UA to M-AC to DP-UA was not. The cycles in each of the graphs are enumerated in Table 3. Examination of these data shows that planning was less systematic in the crossing scenario than in the sequencing scenario, especially in the picture building phase (seven cycles in the crossing graph, but only two in the sequencing graph).

The more predictable sequencing problems resulted in a greater degree of systematic planning, especially in the initial picture building phase. In addition, for the sequencing problem, the management of planning progressed from bottom-up in the initial picture building phase to top-down in the plan development phase. However, this was not the case for the crossing problem. The crossing problem be-

Table 2

Influence Median (values in parentheses).

Phase	Influence Median
Sequencing, Picture building	CD-DAT (1.1)
Crossing, Picture building	CD-DAT (1.3)
Sequencing, Plan development	DP-UA (1.3)
Crossing, Plan development	CD-DAT (1.3) & DP-UA (1.5)
Sequencing, Plan maintenance	DP-UA (1.1)
Crossing, Plan maintenance	CD-DAT (1.4) & DP-UA (1.6)
Sequencing, Replan	DP-UA (1.1) [a]
Sequencing, Replan maintenance	DP-UA (1.2)

[a] This median is not based on all 10 nodes: Not all nodes were reachable because a particular transition never occurred.

gan with bottom-up management but progressed to bi-directional management shared between CD-DAT and DP-UA in the plan development phase.

Plan maintenance

Once the plan was developed and conveyed to the tactician for implementation, the scenario ran for 10 minutes. The planners' task now became maintaining the plan they had created, modifying or reworking it as new aircraft entered the airspace. This was called the plan maintenance phase because the planners appeared to maintain the plan by fine-tuning it, finalizing prior indeterminate decisions (see Dougherty et al., 1999) rather than jettisoning the existing plan as the situation changed. This was possible because planners made definitive decisions about some aircraft (deciding which aircraft would be 1st, 2nd, and 3rd in line to Dallas/Fort Worth), but indeterminate decisions about other aircraft. For example, they would report the group of three aircraft that would be next, but would not yet commit to which would be 4th, 5th, or 6th. Presumably, there was a more optimal time to do that, a few minutes hence, once the uncertainty in the situation had been reduced. Planners were not deferring the entire decision to later (e.g., Hirst & Schweitzer, 1990), just aspects of it. Simultaneously making definitive decisions about some aircraft and indeterminate decisions about others is related to what Kuipers et al. (1988) called embedded planning operators: Physicians stated that, at a certain point in the future, there would be enough information to formulate an appropriate plan. Entin, Needalman, Mikaelian, and Tenney (1988) described a related idea called hedging. They had subjects conduct multiple- or single-option planning in an air-land defense scenario. The single-option subjects assigned and positioned forces in such a way to prepare for two or more future possibilities.

These are all responses on the part of the operator to uncertainty in the environment. The difference lies in the specificity of the alternative plans. A plan incorporating hedging behavior results in the construction of definite, alternative plans. The planners in our study, on the other hand, appeared to consider only one plan, although at any particular moment, aspects of it were not considered in much detail. The advantage of a plan that is simultaneously definite and indefinite was that it could be constantly adjusted and modified to fit the evolving situation. It need not be replaced with a new plan (unless something drastic happened). Note that these ideas are similar to the successive refinement ideas of Sacerdoti (1975). His computer program, NOAH, formulated problems as high-level goals, expanded each goal into subgoals, but maintained indeterminacy as long as possible.

The rolling blocks of 10 verbalizations for the plan maintenance phase are shown in Figure 3. In the sequencing scenarios (bottom panel), a similar, although much less extreme version of the tradeoff pattern observed in Figure 1 was seen for data collection (CD) and plan development (DP). As the scenario began, new aircraft entered the sector, which necessitated additional data collection. These aircraft then needed to be integrated into the plan. Monitoring (M) increased over the course of this phase; problem (P) verbalizations were rare and constant throughout. In the crossing scenario, CD and DP were fairly constant throughout this phase. Monitoring was more frequent than in the sequencing scenarios, sometimes even as frequent as CD and DP. Problem verbalizations were again rare.

The median of the plan maintenance phase for the sequencing graph was DP-UA, signaling that this phase was driven from the top-down. There were no cycles in this graph, signaling that the flow of activity was well orchestrated by DP-UA. A plan was in place, and changes to the situation were incorporated into the normal flow of behavior without prolonged interruption to that flow. In contrast, the crossing scenario reflected a balance between CD-DAT and DP-UA (these two nodes were the medians), and planning activities were much less systematic (9 cycles).

Replan and replan maintenance

After 10 minutes of running the final scenario, which was always one of the sequencing scenarios, the scenario was paused and the planner was handed a note that stated that an accident had occurred at Dallas/Fort Worth and that one half of the Dallas/Fort Worth arrivals had to be rerouted to Oklahoma City and one half had to be rerouted to Houston. This forced the planners to make a new plan. After completing the new plan and conveying it to the tactician, the scenario resumed and continued for an additional 10 minutes. The period in which the scenario was paused was called the *replan phase*; the subsequent 10 minutes was called the *replan maintenance phase*.

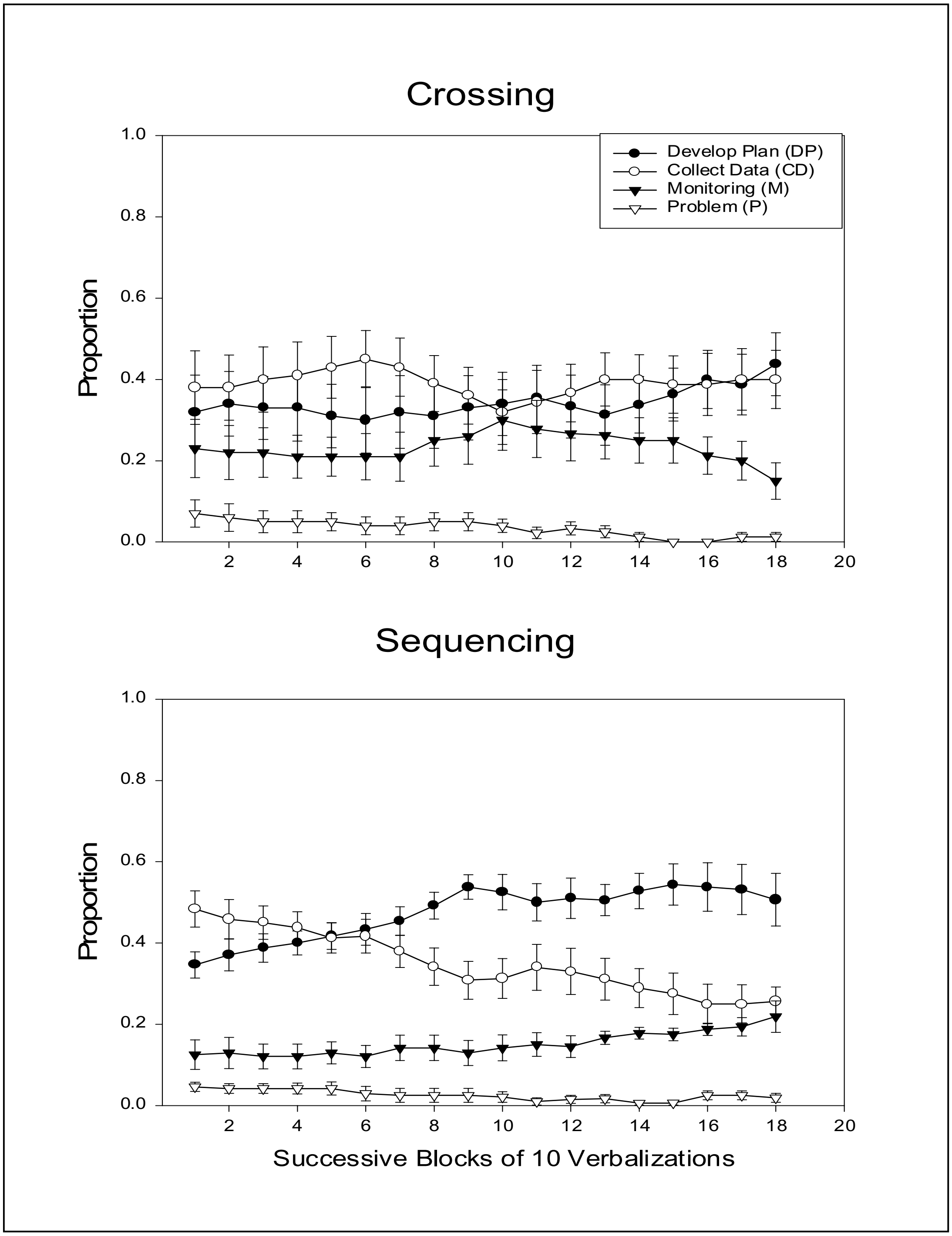

Figure 3. Proportion of verbalizations for the four major categories as a function of rolling blocks of ten verbalizations. Crossing scenario on the top and Sequencing scenario on the bottom. Error bars show one standard error of the mean. Data are from the 10-minute period during which the scenario was running.

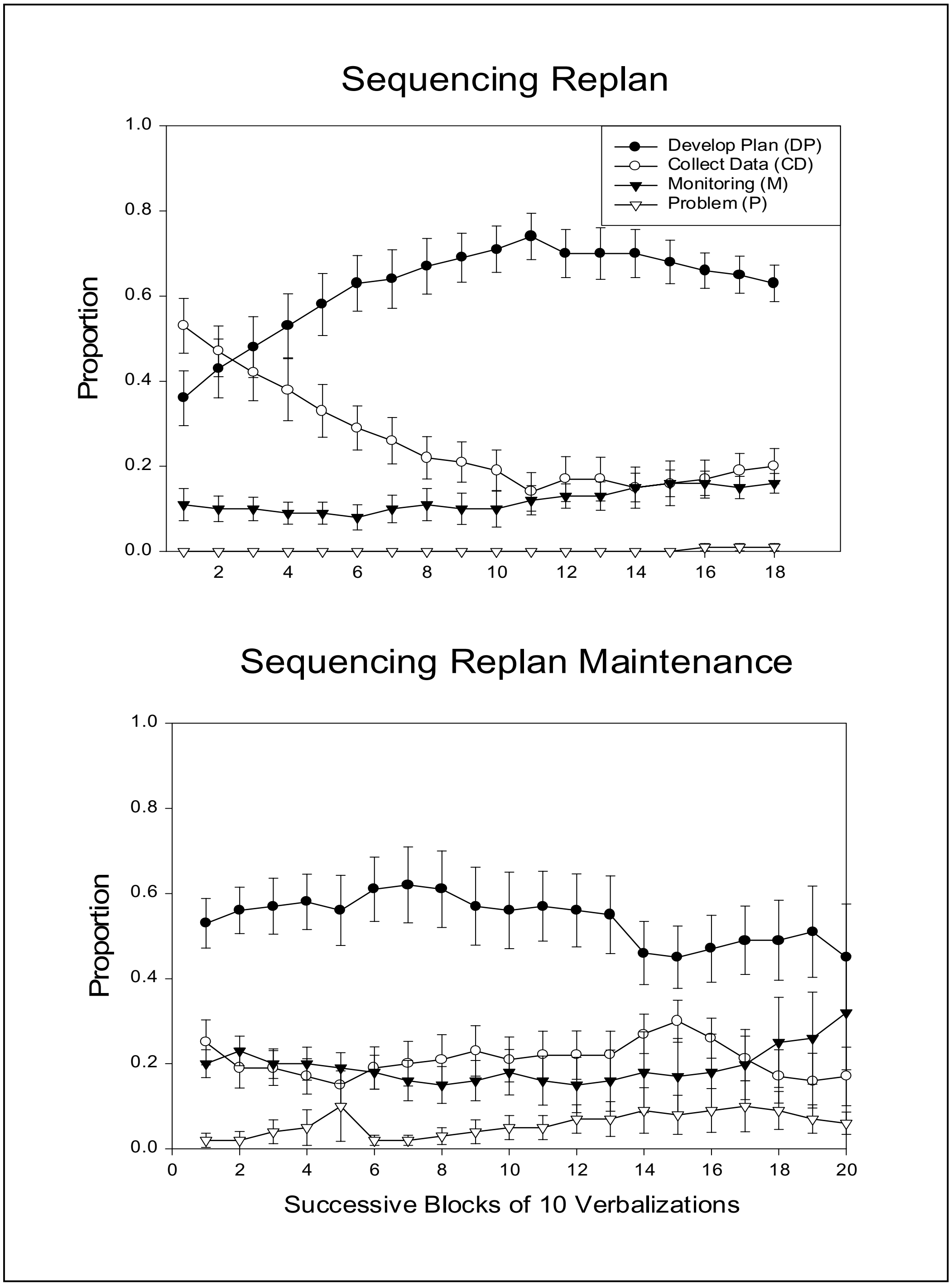

Figure 4. Proportion of verbalizations for the four major categories as a function of rolling blocks of ten verbalizations (error bars show one standard error of the mean): The Sequencing Replan phase is on the top and the Replan Maintenance phase is on the bottom. The scenario was paused during the former, which immediately followed informing the planner about the closing of Dallas/Fort Worth. The latter reflects the 10-minute period during which the scenario was running.

The rolling blocks of 10 verbalizations for these two phases are shown Figure 4. The replan phase is shown in the top panel, the replan maintenance phase in the bottom panel. The replan phase showed a pattern similar to the sequencing scenario in Figure 1. An initial period of data collection, although much briefer than before, was followed by a period of plan development. During the data collection portion of the replan phase, the controller typically checked an aircraft's current location to see if it was easier to take it to Oklahoma City or Houston. The data collection period was not extensive because the planner already knew a lot about these aircraft; they were not new to the scenario as was the case during the initial picture building phase. Once this information was collected, it was time to develop a plan for how to create two new sequences from the one that had been going to Dallas/Fort Worth. The replan maintenance phase was similar to the second half of the plan maintenance phase in the sequencing problem. Develop plan was most frequent and fairly constant; collect data was less frequent but also fairly constant.

No problem verbalizations occurred during the sequencing replan phase, which meant that many links were tied (with the minimum, i.e. zero, transition weight). Because every node must be reachable in a graph, the large number of ties resulted in Pathfinder creating a very large number of links. Consequently, to better reveal the underlying structure, transitions that occurred less than five times in a thousand were eliminated. This greatly simplified the sequencing replan graph (from 54 to 14 links) but did little to the sequencing replan maintenance graph (from 24 to 19 links). Like the sequencing plan maintenance graph in Figure 2, the replan and replan maintenance graphs can be characterized as top-down (median of DP-UA) with a high degree of plan systematicity (0 and 3 cycles).

The replan manipulation provided a within-participant replication of the plan development process. The participants still went through a data collection phase (albeit much briefer) when forced to develop a new plan, despite the fact that they knew a lot about the aircraft already. This suggested that collecting data (or at the very least verification of that data) prior to any major planning activity was the modus operandi of our expert controller participants.

CONCLUSIONS

The present study revealed that, in the air traffic control domain, the direction of plan management depended on the phase of the planning process. In the more predictable sequencing problems, planning began with a bottom-up picture building phase. Completion of the picture building phase was signaled by a sharp transition to a top-down plan development phase. It appears that once a picture of the situation was built, a plan was put into place that guided subsequent behavior. The results were different in the less-predictable crossing problems. Although planning began as a bottom-up process, the next step was a plan development and a plan maintenance phase characterized by equivalent bottom-up and top-down contributions. This might be because the crossing problem did not allow for the development of a single plan to direct activities. Instead, a brief period of planning needed to be followed by a brief period of data collection in preparation for the development of the next mini-plan. A more case-based approach to planning (e.g., Hammond, 1990) would result in a similar interpretation of these data. According to this approach, the prominence of unconditional plan development (DP-UA) signaled that a particular script or specific prior case had been chosen and was coordinating ongoing planning activities. On the other hand, when the direction of management was bi-directional (CD-DAT and DP-UA were equally prominent), it may signal that no single prior case was sufficient to solve the problem.

The degree of plan systematicity depended on the degree of predictability of the problem being solved. The more-predictable sequencing problems resulted in more systematic planning and the less-predictable crossing problem resulted in less systematic planning (as measured by number of cycles). A greater degree of predictability can allow the operator to know what information is important and what information can be ignored. This makes problem states easier to anticipate. It is more difficult to anticipate problem states in a less-predictable environment, and the failure to anticipate problem states requires that an operator frequently adapt to the situation, resulting in less systematic planning.

It is not only unpredictable environments that make it difficult to anticipate problem states. Hayes-Roth and Hayes-Roth (1979) found evidence for unsystematic plan organization in their errand-running task. However, an errand-running task could hardly be considered more unpredictable than the environment facing a pilot or controller (e.g., city buildings do not continuously shift locations). However, the errand subjects were probably insufficiently motivated to gather the information necessary to make problem states known. After all, there was no cost to not gathering the information, contrary to the demands placed on a pilot or controller. Planning errands should move toward being more systematic if expert errand-planners (e.g., FedEx drivers, taxi dispatchers) were highly motivated to succeed.

The results of the present study have implications for theories of planning and for the development of computer interfaces to aid planning. Two aspects of the results are important for theories of planning. First, theories of planning must consider how planning behavior changes over time. Planning is a dynamic process that is responsive to the evolution of the planners understanding of the problem. Theory must explain why these changes are taking place, and what factors are causing the changes. Knowledge of these factors will be important to incorporate into the training of future experts. Second, theories of planning must address the interplay between the planner and the environment in which the planning is taking place. Although issues of environmental predictability are central to some theories of decision-making, these issues have been neglected by theories of planning.

The results also have implications for the construction of planning interfaces. A couple of examples specific to the air traffic control domain will be discussed. Because controllers spend considerable time collecting information prior to plan development, planning interfaces should not just facilitate planning, they should also facilitate picture building (situation assessment). As it stands now, controllers must consult flight progress strips or a Computer Readout Display to find information about destination and routing. One way to enhance data collection is to incorporate a categorization tool that allows for the identification of groups of aircraft (e.g., those going to the same destination, or at the same altitude).

Canning et al. (1999) developed an interface of this type. Aircraft were categorized into color-coded groups according to a user-specified criterion (e.g., all the Dallas/Fort Worth arrivals appeared in orange). Because each group had a unique color on the radar display, controllers only needed to inspect the radar to determine which aircraft belonged together. Consequently, a simple categorization tool circumvented the arduous task of reviewing the strips for this single piece of information. This was in line with a philosophy of automation that keeps the operator in the loop (human-centered automation, e.g., Billings, 1996). A goal of this philosophy is to develop tools that enhance performance by restructuring the information and speeding access to critical information, not by outsourcing the cognitive ability of the operator to the tools. For a similar philosophy applied to air traffic control see the PHARE Highly Interactive Problem Solver developed at Eurocontrol (Meckoff & Gibbs, 1994).

These results also have implications for researchers developing tools that outsource some of a controller's planning ability (e.g., CTAS—Vivona et al., 1996; URET—Brudnicki & McFarland, 1997). In this case, better theories of planning can help a developer decide which tasks should be outsourced and when. For example, planning and picture building tools are equally important in less-predictable problems. However, in more-predictable problems, planning tools may get in the way if an understanding of the situation is not already in place. Perhaps tools could be developed that ensured that a sufficient understanding of the situation was achieved before planning could commence. The scarcity of problem verbalizations might indicate that little benefit would be gained by providing planners with tools that detect problems (although this could be very useful to a tactician). Different types of tools might be needed in environments with differing levels of predictability. This type of environment-specific tool use is evident in some business contexts. For example, the planning tools preferred by business student participants to solve various case studies differed as a function of environmental volatility, level of planning (tactical or strategic), and factors related to the organization (Hartman, White, & Crino, 1986).

SUMMARY

It was possible that more planning took place in the present experiment than normally takes place in the field. That was by design (e.g., the scenario began paused) and was the target of our study. However, we do not think that the experimental setting altered the quality of the planning that took place. These were the types of problems that controllers have to solve every day, using the same equipment they have trained on for years. We also do not think that the distributed verbal protocol method significantly affected the quality of planning. All the participants easily understood the division between planner and tactician; it was a natural division for them. Furthermore, the controllers task is naturally verbal, so having them speak aloud did not change this feature of the task – we only enticed them to be more explicit with respect to planning.

The present study revealed two factors that are important to planning in the air traffic control domain: 1) the phase of planning, and 2) the predictability of the environment. Plan management depended on the phase of the planning process. For the more-predictable problems, planning began with a bottom-up picture building phase followed by a top-down plan development phase. The less-predictable problem began as a bottom-up process, but management of the subsequent phases was characterized by equivalent bottom-up and top-down contributions. Degree of predictability affected plan systematicity. More systematic planning took place for predictable problems and less systematic planning for less predictable problems. Future research should explore both the generality of these findings to other domains and the extent to which the findings generalize to field controllers working familiar sectors.

REFERENCES

Adams, M.J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. Special Issue: Situation awareness. *Human Factors*, *37*, 85-104.

Bainbridge, L. (1997). The change in concepts needed to account for human behavior in complex dynamic tasks, *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, *27*, 351-359.

Billings, C.E. (1996). Human-centered aviation automation: Principles and guidelines. *NASA Technical Memorandum 110381*. Moffett Field, CA

Brudnicki, D.J., & McFarland, A.L. (1997). User Request Evaluation Tool (URET) Conflict Probe Performance and Benefits Assessment, *MP97W112*. McLean, VA: The MITRE Corporation.

Byrne, R.W. (1977). Planning meals: Problem-solving on a real data-base. *Cognition*, *5*, 287-332.

Canning, J.M., Johansson, J., Gronlund, S.D., Dougherty, M.R.P., & Mills, S.H. (1999). Effects of a novel interface design on strategic planning by en route controllers. *Proceedings of the 10th International Symposium on Aviation Psychology*, Columbus, OH.

Carlson, L.S., Rhodes, L.R., & Cullen, M.G. (1996). Effects of unstructured routes on en route controllers' work activities and operational environment, *MITRE Technical Report 96W19*. McLean, VA: The MITRE Corporation.

Dougherty, M.R.P., Gronlund, S.D., Durso, F.T., Canning, J. & Mills, S.H. (1999). Plan generation in air traffic control. *Proceedings of the 10th International Symposium on Aviation Psychology*, Columbus, OH.

Durso, F.T., & Gronlund, S.D. (1999). Situation awareness. In F.T. Durso, R. Nickerson, R. Schvaneveldt, S. Dumais, M. Chi, & S. Lindsay (Eds.), *The Handbook of Applied Cognition* (pp. 283-314). New York: Wiley.

Durso, F.T., Rea, C.B., & Dayton, T. (1994). Graph-theoretic confirmation of restructuring during insight. *Psychological Science*, *5*, 94-98.

Durso, F.T., Truitt, T.R., Hackworth, C.A., Ohrt, D.D., Hamic, J.M., & Manning, C.A. (1996). Factors characterizing en route operational errors: Do they tell us anything about situation awareness? In D.J. Garland & M.R. Endsley (Eds.), *Proceedings of the International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp.189-196). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Entin, E.E., Needalman, A., Mikaelian, D., & Tenney, R.R. (1988). *Experiment II report: The effects of option planning and battle workload on command and control* (TR-388-1). Burlington, MA: ALPHATECH, Inc.

Federal Aviation Administration (1995). *Air Traffic Service Plan: 1995-2000*. Washington, DC: Author.

Gigerenzer, G., Todd, P.M., & The ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Goldsmith, T.E., & Davenport, D.M. (1990). Assessing structural similarity of graphs. In Schvaneveldt, R. W. (Ed.), *Pathfinder associative networks: Studies in knowledge organization* (pp. 75-87). Norwood, NJ: Ablex Publishing.

Gronlund, S.D., Ohrt, D.D., Dougherty, M.R.P., Perry, J.L., & Manning, C.A. (1998). Role of memory in air traffic control. *Journal of Experimental Psychology: Applied*, *4*, 263-280.

Hammond, K.J. (1990). Case-based planning: A framework for planning from experience. *Cognitive Science*, *14*, 385-443.

Hammond, K.R., Hamm, R.M., Grassia, J., & Pearson, T. (1997). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. In W.M. Goldstein & R.M. Hogarth's (Eds.) *Research on Judgment and Decision Making: Currents, Connections, and Controversies (pp.144-180)*. New York: Cambridge University Press.

Hartman, S.J., White, M.C., & Crino, M.D. (1986). Environmental volatility, system adaptation, planning requirements, and information-processing strategies: An integrative model. *Decision Sciences*, *17*, 454-474.

Hayes-Roth, B., & Hayes-Roth, F. (1979). A cognitive model of planning, *Cognitive Science*, *3*, 275-310.

Hirst, E., & Schweitzer, M. (1990). Electric-utility resource planning and decision-making: The importance of uncertainty. *Risk Analysis*, *10*, 137-146.

Hutton, R.J.B., Olszewski, R., Thordsen, M.L., & Kaempf, G.L. (1997). En route air traffic controller decision making model and decision maker vulnerabilities, (Contract for NYMA Inc., Egg Harbor Township, NJ). Fairborn, OH: Klein Associates, Inc.

Johannsen, G., & Rouse, W.B. (1983). Studies of planning behavior of aircraft pilots in normal, abnormal, and emergency situations. *IEEE Transactions on Systems, Man, & Cybernetics*, *SMC-13*, 267-278.

Johnson, S.C. (1967). Hierarchical clustering schemes. *Psychometrika*, *32*, 241-254.

Kellogg, R.T. (1994). *The Psychology of Writing*. New York: Oxford University Press.

Kirlik, A., & Bisantz, A.M. (1999). Cognition in human-machine systems: Experiential and environmental aspects of adaptation. In Hancock, Peter A. (Ed.), *Human performance and ergonomics,* (pp. 47-68). San Diego, CA: Academic Press.

Klein, G.A. (1989). Recognition-primed decisions. In W. B. Rouse (Ed.), *Advances in man-machine systems research* (Vol. 5, pp. 47-92). Greenwich, CT: JAI.

Kuipers, B., Moskowitz, A.J., & Kassirer, J.P. (1988). Critical decisions under uncertainty: Representation and structure. *Cognitive Science*, *12*, 177-210.

Meckoff, C., & Gibbs, P. (1994). PHARE Highly Interactive Problem Solver, *Eurocontrol report 273*, Brussels, Belgium.

Moertl, P.M., Canning, J.M., Johansson, J., Gronlund, S.D., Dougherty, M.R.P., & Mills, S.H. (2000). Workload and Performance in FOPA: A Strategic Planning Interface for Air Traffic Control. *Proceedings of the 44th Annual Meeting of the Human Factors and Ergonomics Society*, 3-65 – 3-68.

Mumford, M.D., Shultz, R., & Van Doorn, J.A. (in press). Performance in planning: Processes, requirements and errors. *Review of General Psychology*.

Payne, J.W. (1994). Thinking aloud: Insights into information processing. *Psychological Science*, *5*, 245-248.

Pew, R.W., & Mavor, A.S. (Eds.) (1998). *Modeling Human and Organizational Behavior: Application to Military Simulations*. Washington, DC: National Academy Press.

Pietras, C.M. & Coury, B.G. (1991). Cognitive models of planning in the design of project management systems. *Proceedings of the Human Factors Society*, 416-420.

RTCA (1995). *Final report of the RTCA Task Force 3: Free flight implementation*. Washington, DC: RTCA Incorporated.

Sacerdoti, E.D. (1975). A structure for plans and behavior. *Technical Note 109*, Stanford Research Institute, Menlo Park, CA.

Schvaneveldt, R.W. (1990). *Pathfinder associative networks: Studies in knowledge organization*. Norwood, NJ: Ablex Publishing.

Schvaneveldt, R.W., Durso, F.T., & Dearholt, D.W. (1989). Network structures in proximity data. *Psychology of Learning and Motivation*, *24*, 249-284.

Shepard, R.N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance: I *Psychometrika*, *27*, 125-140.

Stewart, S.I., & Vogt, C.A. (1999). A case-based approach to understanding vacation planning. *Leisure Sciences*, *21*, 79-95.

Velleman, P.F., & Hoaglin, D.C. (1981). *Applications, basics, and computing of exploratory data analysis*. Boston, MA: Duxbury Press.

Vivona, R.A., Ballin, M.G., Green, S.M., Bach, R.E, & McNally, B.D. (1996). A system concept for facilitating user preferences in en route airspace, *NASA Technical Memorandum 4763*.

Vortac, O.U., Edwards, M.B., & Manning, C.A. (1994). Sequences of actions for individual and teams of air traffic controllers. *Human-Computer Interaction*, *9*, 319-343.

Wilson, T.D. (1994). The proper protocol: Validity and completeness of verbal reports. *Psychological Science*, *5*, 249-251.

Xiao, Y., Milgram, P., & Doyle, D.J. (1997). Planning behavior and its functional role in interactions with complex systems. *IEEE Transactions on systems, man, and cybernetics—Part A: Systems and humans*, *27*, 313-324.