IN-64-CR
8362
p. 26

Design of a Cooperative Problem-Solving System for En-Route Flight Planning: An Empirical Evaluation

CHARLES LAYTON, *Galaxy Scientific Corporation, Atlanta, Georgia*, PHILIP J. SMITH,[1] *Ohio State University, Columbus, Ohio, and* C. ELAINE MC COY, *Ohio University, Athens, Ohio*

Both optimization techniques and expert systems technologies are popular approaches for developing tools to assist in complex problem-solving tasks. Because of the underlying complexity of many such tasks, however, the models of the world implicitly or explicitly embedded in such tools are often incomplete and the problem-solving methods fallible. The result can be "brittleness" in situations that were not anticipated by the system designers. To deal with this weakness, it has been suggested that "cooperative" rather than "automated" problem-solving systems be designed. Such cooperative systems are proposed to explicitly enhance the collaboration of the person (or a group of people) and the computer system. This study evaluates the impact of alternative design concepts on the performance of 30 airline pilots interacting with such a cooperative system designed to support en-route flight planning. The results clearly demonstrate that different system design concepts can strongly influence the cognitive processes and resultant performances of users. Based on think-aloud protocols, cognitive models are proposed to account for how features of the computer system interacted with specific types of scenarios to influence exploration and decision making by the pilots. The results are then used to develop recommendations for guiding the design of cooperative systems.

INTRODUCTION

Three alternative designs for a cooperative problem-solving system (Robertson, Zachery, and Black, 1990) were empirically evaluated in this study. All three designs provided support for the task of en-route planning for commercial aviation flights. The designs differed in terms of the timing and degree of assistance provided by the computer. Our goals were to gain a better understanding of how people perform adaptive planning tasks, to increase our understanding of how alternative system designs influence the cognitive processes of users during such planning tasks, and to develop recommendations to guide in the design of advanced tools to support pilots

[1] Requests for reprints should be sent to Philip J. Smith, Cognitive Systems Engineering Laboratory, 210 Baker Systems, 1971 Neil Ave., Columbus, OH 43210.

and dispatchers in their flight-planning activities. Of particular interest was the question: How does the introduction of computer-generated alternative plans change the exploration of data and alternative routes and the evaluation of these alternative routes by the human planner?

CONTEXT

En-route flight planning (Cohen, Leddo, and Tolcott, 1989; Johannsen and Rouse, 1983; Rudolf, Homoki, and Sexton, 1990; Sorensen, Waters, and Patmore, 1983) involves the modification of the aircraft route of flight (flight plan) of an airborne aircraft in response to problems with weather, air traffic, medical emergencies, mechanical failures, and so on. The flight crew, air traffic controllers, and airline company dispatchers all play important roles in this planning process.

The flight plan stipulates at what altitude and heading the aircraft will fly during various phases of the flight and what route the aircraft will take. The route determines the weather and air traffic that will be encountered along the way, which affects speed, safety, fuel efficiency, passenger comfort, and arrival time.

The planner, then, is concerned with getting from a given origin to a given destination in a timely and cost-effective fashion while maintaining flight safety and passenger comfort. The planner must consider what routes to take (these routes consist of navigational fixes and jet routes, the so-called highways in the sky that connect the navigational fixes), at what altitudes to fly, what weather to avoid (including winds, thunderstorms, freezing rain, and turbulence), and the ever-changing characteristics of the aircraft (for example, the decrease in weight as fuel is consumed).

The initial flight plan is rarely followed exactly, because of unforeseen events en route.

Indeed, minor changes in flight plans are frequently made, and major changes are common. These amendments to the original plan result from the dynamic, unpredictable nature of the "world" in which the plans are executed. Weather patterns do not always develop as predicted, resulting in unexpected areas of turbulence, less favorable tailwinds, or storms that must be avoided. Air traffic congestion may delay takeoff or restrict the aircraft to altitudes that are lower than planned. Airport or runway closures can cause major disruptions not only for one aircraft but also for everyone planning to land at that airport. Mechanical failures, medical emergencies, or other critical problems may force the aircraft to divert to an unplanned airport.

In short, en-route flight planning is a large and complex problem. Multiple goals must be considered in a highly stochastic environment in which multiple plans must be coordinated (Hayes-Roth and Hayes-Roth, 1979; Hoc, 1988; Miller, Galanter, and Pribram, 1960; Sacerdoti, 1974; Schank and Abelson, 1977; Stefik, 1981; Suchman, 1987; Wilensky, 1983).

THE FLIGHT-PLANNING TEST BED: DESIGN FEATURES

The Flight-Planning Test Bed was developed to test several cooperative planning system design concepts (Coombs and Alty, 1987; Lehner and Zirk, 1987; Shute and Smith, 1993; Thierauf, 1988). This design was developed following an extensive cognitive task analysis (Smith, McCoy, Layton, and Bihari, 1992). The basic flight-planning system performs a number of functions in response to input from a human operator. The system allows the user (either a pilot or a dispatcher) to develop and display up to four flight plans in conjunction with weather information and

to obtain feedback in terms of such flight parameters as fuel, time, and distance. The weather information consists of both graphic depictions and oral descriptions and can be displayed at several altitudes. The displays show the entire flight path, thus emphasizing global solutions to problems. In addition, the user can manipulate the display time to see the relationship between the weather information and the aircraft's position. The system computes the optimal altitude profile to minimize fuel consumption, arrival times at navigational fixes, and fuel remaining at those fixes, based on wind components. It will also determine these flight parameters given a user-selected altitude profile.

The basic system runs on a Macintosh IIfx with two color monitors. The features and functions on each monitor are discussed in turn.

Left Monitor

The displays and controls on the left monitor are shown in Figure 1. (In all of the figures that depict system displays, some of the information loses salience as printed here in black and white instead of color.)

The primary feature on the left monitor is a map display. This display depicts the continental United States, the aircraft position, and planned routes. Several pieces of information can be overlaid on this map, including weather information (overlays of composite cloud and composite radar charts, fronts, and cloud cover, radar, and winds at specific altitudes), navigational fixes, and jet routes (see Figure 2).

All weather information is available for two display times: the current time and a 1-h forecast. When a forecast is displayed, the aircraft is displayed in its predicted position
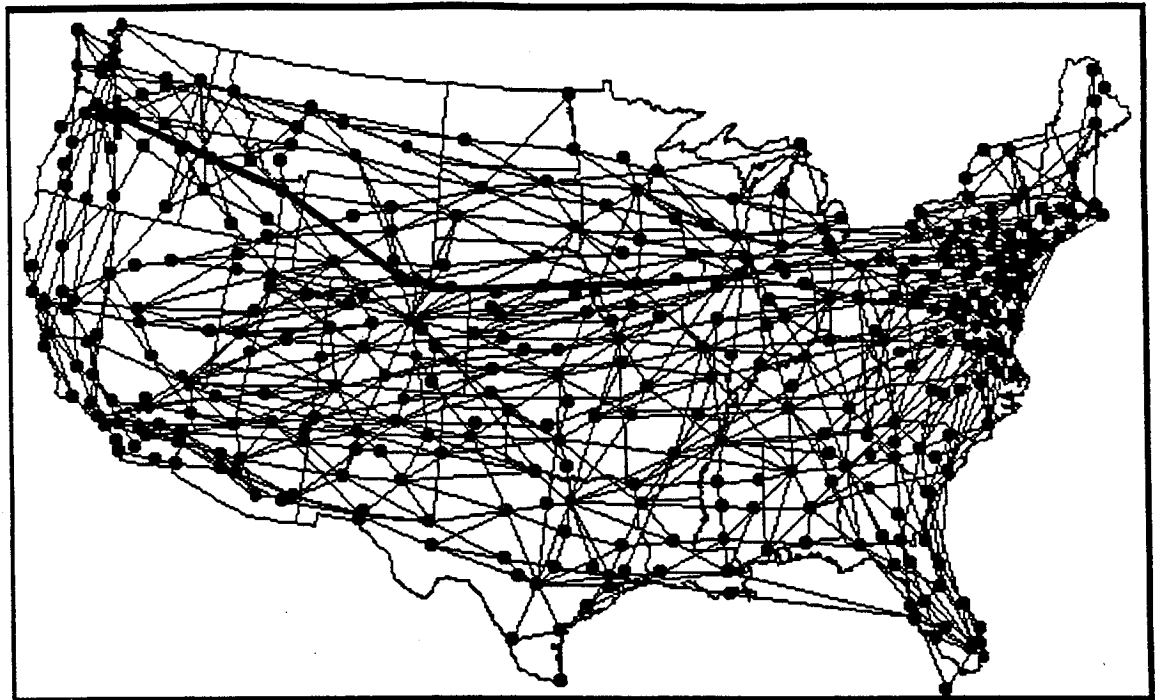

Figure 1. *Left monitor display: original route (the middle route) and routes explored by sample subjects in Case 1.*
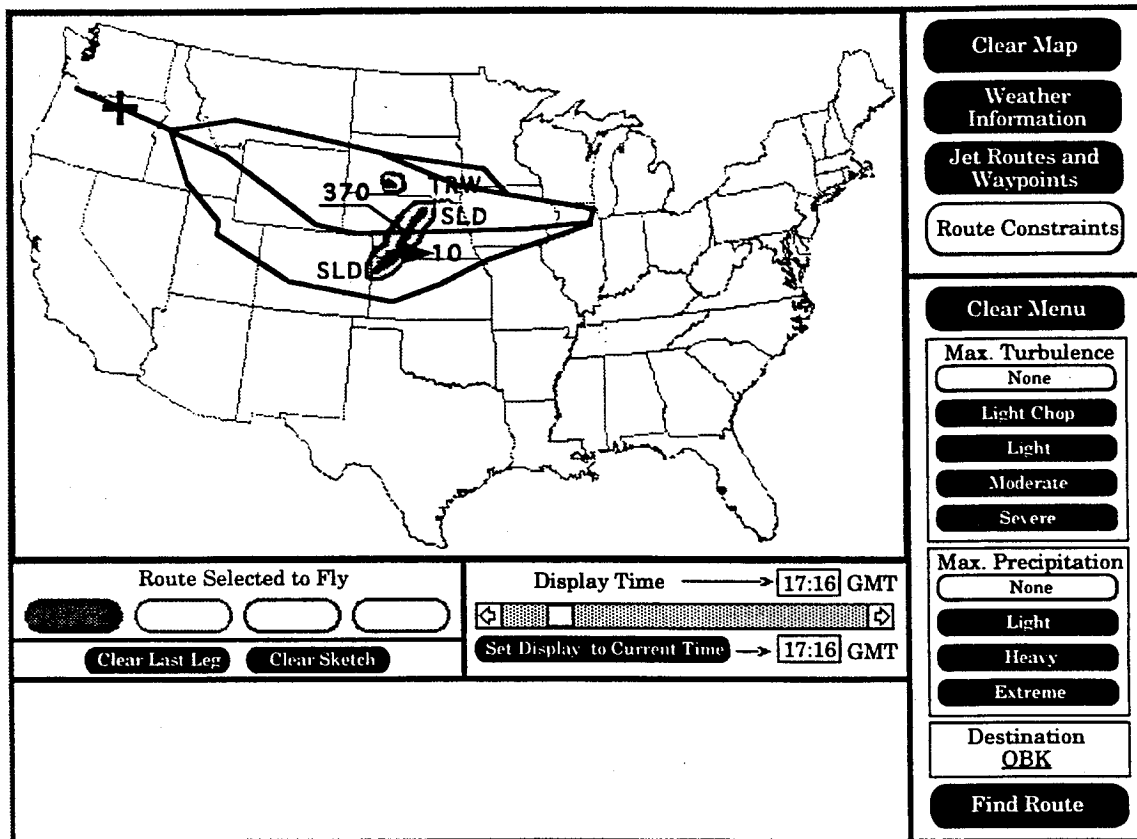
Figure 2. *High-altitude jet routes and navigational fixes.*

(on each route) at the forecast time, as well as in its current position. The user can also zoom in on a region of the map, which replaces the map of the continental United States with a magnification of an area surrounding a user-selected point.

Right Monitor

The right monitor displays and controls are shown in Figure 3. It displays a flight log of a route. This flight log is essentially a spreadsheet that depicts each segment of the route (i.e., all of the navigational fixes and jet routes that make up the route), as well as information pertinent to those segments. The flight log also graphically displays the planned altitudes and the least-fuel-consumption altitudes for that route. Finally, the flight log displays weather information that is pertinent to the route. For example, turbulence information is displayed by default, but the operator can also select wind information. The turbulence information that is presented is a one-word summary of the maximum turbulence on a given flight segment. The operator can get a more detailed description of that information (available pilot reports, or "pireps") by selecting (clicking on) the one-word summary.

The other display on this monitor (at the bottom of the screen) shows important flight parameters for all four alternative routes
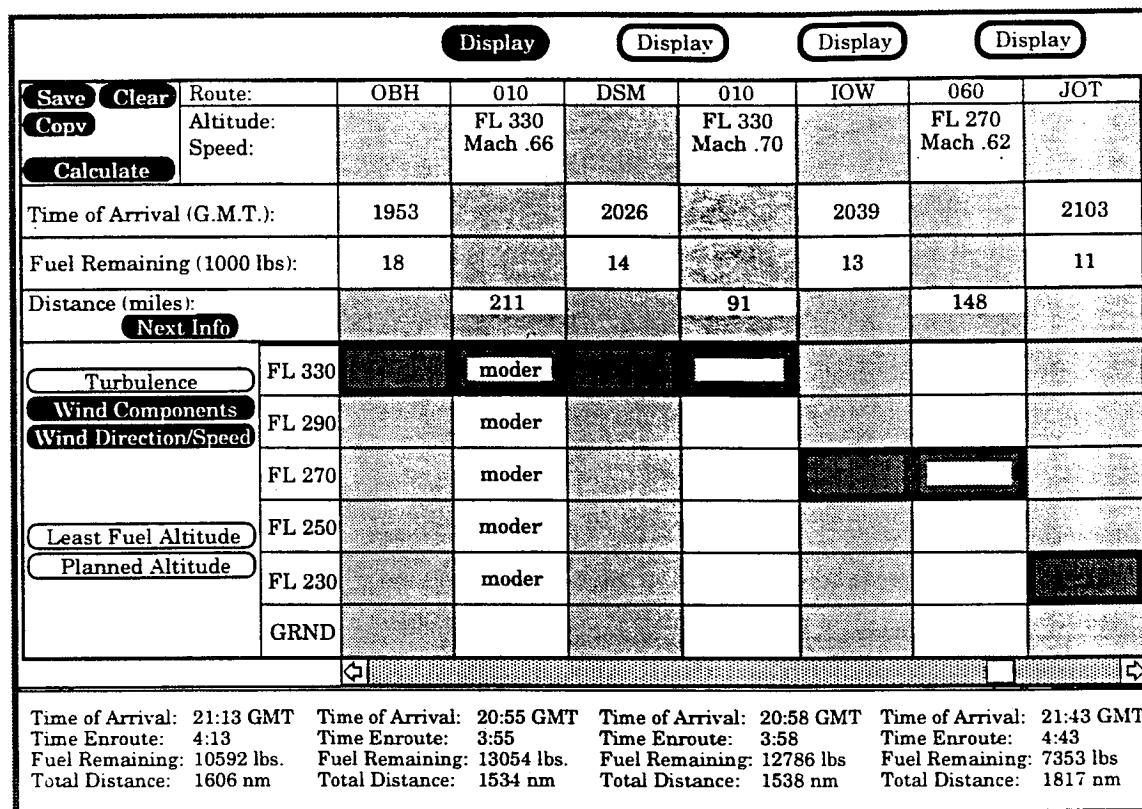
			Display		Display		Display		Display
Save **Clear**	Route:		OBH	010	DSM	010	IOW	060	JOT
Copy	Altitude: Speed:			FL 330 Mach .66		FL 330 Mach .70		FL 270 Mach .62	
Calculate									
Time of Arrival (G.M.T.):			1953		2026		2039		2103
Fuel Remaining (1000 lbs):			18		14		13		11
Distance (miles): **Next Info**				211		91		148	
Turbulence	FL 330			moder					
Wind Components **Wind Direction/Speed**	FL 290			moder					
	FL 270			moder					
Least Fuel Altitude	FL 250			moder					
Planned Altitude	FL 230			moder					
	GRND								

Time of Arrival:	21:13 GMT	Time of Arrival:	20:55 GMT	Time of Arrival:	20:58 GMT	Time of Arrival:	21:43 GMT
Time Enroute:	4:13	Time Enroute:	3:55	Time Enroute:	3:58	Time Enroute:	4:43
Fuel Remaining:	10592 lbs.	Fuel Remaining:	13054 lbs.	Fuel Remaining:	12786 lbs	Fuel Remaining:	7353 lbs
Total Distance:	1606 nm	Total Distance:	1534 nm	Total Distance:	1538 nm	Total Distance:	1817 nm

Figure 3. *Right monitor display.*

upon arrival at the destination. These parameters include time of arrival, time en route, fuel remaining, and total distance. This display allows users to compare the "bottom line" for each route.

Important Features

The design principles underlying the Flight-Planning Test Bed as a cooperative planning system are discussed in detail in Smith et al. (1992). Five of the most significant objectives, however, are as follows:

(1) To provide tools that allow cooperative planning at different levels of abstraction (inspired by the work of Hayes-Roth and Hayes-Roth, 1979; Sacerdoti, 1974; Shute and Smith, 1993; Suchman, 1987)
(2) To provide the human planner with data dis-

plays and representations to support plan generation and evaluation at these different levels of abstraction
(3) To provide interfaces to the available support tools that allow the person to easily communicate desired trade-offs among goals
(4) To provide tools that help the person predict the outcomes of various plans (Coombs and Alty, 1987)
(5) To incorporate a graphical interface that allows the person to view and explore alternative plans in the context of the relevant data (i.e., weather displays)

In the next section we describe an empirical study to assess some of these design considerations.

METHODS

In the following study, the Flight-Planning Test Bed was used to study the effects of

83

different design features on cooperative problem-solving performance. Briefly, each of the 30 subjects (professional airline pilots) was asked to use one of three alternative system designs (10 subjects per condition). Each subject was trained in the use of that particular system and given four cases to solve.

System Designs

Three different en-route flight-planning support systems were designed that represented variations on the levels and timing of support provided by the computer. These variations on the system design represented the independent variable studied in this experiment.

Sketching-only system. The sketching-only system allowed the human planner to sketch proposed flight paths on a map display while the computer filled in lower-level details (such as fuel remaining, time of arrival, and recommended altitudes) using an optimization program. In this version the person was responsible for proposing the alternate paths, whereas the computer was responsible for providing feedback on those solutions.

The sketching of routes was carried out by displaying the jet routes and navigational fixes and selecting (clicking on) each navigational fix through which the pilot wanted the airplane to pass. This placed a slight restriction on planning because vectoring can normally be requested to fly direct routes from one point to another. However, this approach allowed the planner to develop general solutions with the understanding that these solutions were not necessarily the exact routes that would be flown.

Route constraints and sketching system. The route constraints and sketching system retained all of the capabilities of the sketching-only system and added another capability: Operators could specify higher-level constraints on the solution they desired and then

ask the computer to find the shortest route that satisfied those constraints. The constraints that could be specified were maximum allowable turbulence, maximum allowable precipitation, and destination. (It is easy to see how this interface design concept could be extended to include other constraints, such as earliest and latest desired arrival times or number of passengers making their connections.)

Users could specify constraints on the solution they desired from the computer. The computer would then recommend alternatives. In addition, by means of the sketching tool, users could explore specific routes themselves.

Automatic route constraints, route constraints, and sketching system. This version took the computer's involvement one step further: The computer automatically suggested a deviation (based on default constraints of no turbulence, no precipitation, and the originally planned destination) as soon as it detected a problem with the original route. This form of tool is akin to an autonomous support system that automatically suggests solutions to detected problems. This system also made available the route constraints tool of the previous system and the sketching tool of the previous two systems.

Underlying all three system designs is the incorporation of tools to support asking "what if" questions. That is, these tools help the operator to ask such questions as, "What type of solution does the computer suggest if I use constraints of light turbulence and moderate precipitation?" or "What happens to my fuel remaining if I deviate north instead of south?" We were interested in whether people used the tools available to them, how the available tools affected the cognitive processes of the person using the system, and how the available tools affected the solutions that person chose.

Subjects

Thirty male commercial airline pilots volunteered to participate in this 3-h study. (Approximately half of that time was spent training the pilot on the system he would be using.) Each pilot was randomly assigned to one of the three alternative system designs. The pilots represented eight major airlines, and had an average of 9300 h of flying experience as commercial pilots and 1800 h of experience in military aircraft. In the results to be presented, no apparent relationships were found between the pilots' performances and their levels or types of flying experience, or with their levels of previous computer experience.

Cases: Characteristics and General Predictions

Following training on the use of the system, each of the subjects was presented with four en-route flight-planning cases in which he was given some preliminary information about the flight (e.g., origin, destination, time of day) and was then told to "decide what the aircraft should do." All of the subjects went through the same four cases in the same order. Whereas the subjects in the sketching-only and route constraints and sketching conditions started each case with only their original route of flight, the subjects in the automatic route constraints, route constraints, and sketching condition were also given an alternate route suggested by the computer based on the default constraints of finding a route that was predicted to avoid all turbulence and precipitation.

Cases 1 through 3 can be characterized as having large *solution spaces;* that is, the number of plausible specific flight paths available to accomplish a particular deviation (such as going north of the storm) was very large. This characteristic was expected to put the subjects in the sketching-only version at a disadvantage in terms of finding fuel and time-efficient alternative routes. It was also expected to cause the sketching-only subjects to develop a larger number of specific flight plans for comparison.

All four cases could be described as having a large *data space*, in the sense that the types and amounts of data available in the different displays were fairly large (though still small by comparison with the types and amounts currently being proposed for commercial systems at several of the major airlines). This characteristic was expected to be most important in Case 2, in which a failure to note the head winds to the south might lead to selection of the less efficient southern deviation by subjects sketching their own solutions, and in Case 3, in which a failure to look carefully at the location of the aircraft in relation to current and forecast weather might contribute to the acceptance of a poor plan.

Case 3 had the further property that the limitations of the computer's knowledge led to "brittle" performance, in which the computer generated a poor suggestion for an alternative route. This brittleness was attributable to the fact that in searching for flight plans, the computer treated forecasts as reality. Uncertainty associated with the forecasts was not considered in the computer's reasoning. This case was included to determine whether subjects in the automatic version would be more likely to be drawn into the computer's "world," consequently failing to use their own knowledge of the uncertainty associated with such a forecast to reject the computer's recommendation.

Finally, Case 4, which has a much smaller solution space, is interesting because it introduces a conflict between a common heuristic used in selecting a flight amendment and the fuel and time efficiency of the alternative routes. Further details on these cases are included in discussions of the results.

RESULTS AND DISCUSSION

Case 1

The following scenario was read to subjects prior to their working on this case:

> It is summer and you're on a flight from Battleground (Portland) to Northbrook. Your dispatcher gave you a southerly route in order to avoid an occluded front. The front has dropped to the south as well, however, and has generated some thunderstorms. Time out was 1700 Zulu and you are five minutes into the flight. Decide what you think the aircraft should do.

In all cases, for subjects in the treatment condition in which the computer automatically suggested a solution on loading the case, the following two lines were added (prior to "Decide what you think . . ."):

> The computer has suggested the orange route as an alternative to the original plan (the green route) based on constraints of no turbulence and no precipitation. You may accept either of these plans or develop your own.

The original route (the middle route), the current aircraft position, and the current composite radar are shown in Figure 1. The radar returns show a solid line of thunderstorms with cell tops at 37 000 ft. (For this experiment the pilots were told the aircraft's maximum altitude was 33 000 ft.) Furthermore, the gap between the two cells was forecast to close. Therefore, a deviation was obviously required. The forecast storm movement was to the east but was very small. To provide a concrete sense of the performances of the subjects, the behaviors of three representative pilots are first summarized. Then summary statistics are provided for the entire group.

Subject S1: sketching-only system. Subject S1 looked at the composite radar and fronts (current and forecast) and concluded, "Going to have to go north or south around it." This pilot then sketched a northern deviation and compared it with the original route, noting that the deviation saved time and fuel and avoided the turbulence. He then sketched a

southern deviation. While sketching the deviation, he inferred that "it could move a little further south [than forecast]," so he adjusted the southern alternative for that contingency. When the route was completed, he looked at the computer's estimates for time and fuel consumption and stated, "That one's quite a bit longer," and decided to deviate north.

Subject C3: route constraints and sketching system. Subject C3 looked at the composite radar and concluded, "I can see right now that what I want to do is come to the north." After also looking at the clouds, he let the computer find a deviation based on constraints of light turbulence and light precipitation. The subject looked at the resultant northern deviation suggested by the computer and stated, "That looks like about what I would have in mind." After checking the data displayed on the national map to make sure that the northern deviation had "no problem with turbulence or precipitation," he compared it with the original route and noted, "The total distance is actually a little less. Fuel left is more, and we'll actually cut time off our flight with this route." He then decided to fly the computer-recommended northern deviation.

Subject A9: automatic route constraints, route constraints, and sketching system. The computer automatically displayed to this pilot a recommendation north of the storm. He began his evaluation by comparing the estimated time and fuel consumption for this suggested route (to the north of the storm) with the performance parameters for the original route. He looked at the composite radar and turbulence on the national map and concluded, "The alternate route certainly looks better to me, and I would stick with that."

Comparisons of sample subjects. Figure 1 shows the routes explored in detail by these sample subjects. The subject in the sketching-only version of the Flight-Planning Test Bed

86

(S1) explored the far northern route and the southern route and elected to take the far northern route. The subject who had access to both the route constraints function and the sketching function (C3) elected to use the route constraints function. He accepted the computer's suggested route (the northern route that deviates tighter to the storm in Figure 1) without exploring any other alternatives in detail. The subject who was automatically shown the computer's suggestion (based on constraints of no turbulence and no precipitation) showed evidence of evaluating a southern deviation at an abstract level but chose to accept the computer's suggestion to the north without exploring any other alternatives in detail. (He discussed the tailwinds along a southerly deviation but did not sketch any such plan.)

The primary difference of interest (as illustrated by these three sample subjects) is the fact that the pilot who had to sketch his own solution selected a more conservative route, staying farther north of the storm. This difference, which was found in many of the subjects using the sketching-only version of the system (to be summarized later) appears to result from the tendency of those pilots who sketched their own solution to more thoroughly consider the uncertainty associated with the forecast. (Additional data pertinent to this result are discussed next.)

Case 1 Summary Statistics

The following is a description of behaviors for all of the subjects in the different treatment conditions for Case 1.

Differences in final routes. In this case all 30 subjects chose to deviate to the north, which is not surprising considering that this class of solutions is more direct than the original routing and is closer to the route that someone would fly if there were no weather problems. Nine of the ten subjects in the route constraints and sketching version and nine of the ten subjects in the automatic route constraints version selected the computer's suggested northern deviation. Only two of the ten subjects in the sketching-only version selected that route. Six of the ten subjects in the sketching-only version selected a more conservative northern deviation. Based on a chi-square test, these differences are significant at $\alpha < 0.004$.

Differences in exploration. Of the 10 subjects using the sketching-only version, 4 explored multiple classes of solutions in detail (i.e., they generated actual flight plans on the screen). None of the 20 subjects using the other two versions explored more than a single class of solutions in detail. (In Case 1 exploring a solution north of the storm was defined as one class of solution, and exploring a solution south of the storm was a second class.) This difference was significant at $\alpha < 0.01$.

Another measure of the amount of exploration is the number of subjects who explored multiple specific solutions in detail (as contrasted with multiple classes of solutions, as summarized earlier). Again, the subjects in the sketching-only version showed evidence of more exploration ($\alpha < 0.014$; 6 in 10 subjects vs. 1 in 10 for both of the other two versions).

Differences in information search. The information that the subjects looked at was also analyzed on the basis of treatment condition. The number of subjects in each condition who looked at fronts, current or radar weather (composite or at altitude), winds, and jet routes is presented in Table 1.

As can be seen from this table, there are no clear differences on Case 1 between groups in information displayed, with the exception that more of the sketching-only and route constraints and sketching subjects looked at the jet routes than did the automatic route constraints, route constraints, and sketching subjects. This latter fact is evidence that

TABLE 1

Information Displayed by Subjects

	Fronts	Radar	Winds	Jet Routes
Case 1				
Sketch	9	9	7	10
Constraints	8	10	5	9
Automatic	9	9	6	5
Case 2				
Sketch	7	10	7	10
Constraints	7	10	3	9
Automatic	8	9.	7	8
Case 3				
Sketch	10	10	5	10
Constraints	6	10	1	7
Automatic	8	10	5	9
Case 4				
Sketch	8	10	6	10
Constraints	7	10	4	8
Automatic	6	10	4	9

many of the automatic route constraints subjects evaluated the suggested route at an abstract level (i.e., they didn't look at specific choices of alternative jet routes to deviate north of the storm). This one difference is significant at $\alpha < 0.013$.

Case 1 Discussion

As described earlier, prior to the experiment we made two predictions that are relevant to these results. We predicted that, in general, the pilots using the automatic version might be less likely than the sketching-only subjects to explore as many alternatives in detail and less likely than the sketching-only subjects to consider the uncertainty associated with weather forecasts, consequently accepting the computer's recommendation without adequate evaluation. The results for Case 1 are consistent with these general predictions. They indicate, for example, that the sketching-only subjects explored more alternatives. Furthermore, the concurrent oral reports indicate that the sketching-

only subjects who deviated farther north (see Figure 1) were indeed considering the uncertainty associated with the forecast, making statements such as, "If the system moved further north and the thunderstorms started to pop up" and "Let's take a look at how much further north we could go."

One way to explain these effects is to say that the pilots in the automatic conditions were overreliant or overtrusting of the system, failing to critically evaluate the situation and the computer-suggested plan. These are rather shallow labels, however, and don't provide much insight into the influence of the system's design on the user's cognitive processes.

The clearest example of such an effect occurred in Case 1 at the point at which the subjects had to decide whether to stay north of the storm, from DPR to RWF, or to begin turning south toward the destination, from DPR to FSD (see Figure 4). Because the system design induced the sketching-only subjects to view the display shown in this figure if they wanted to complete a reasonable northern deviation, it can be concluded that
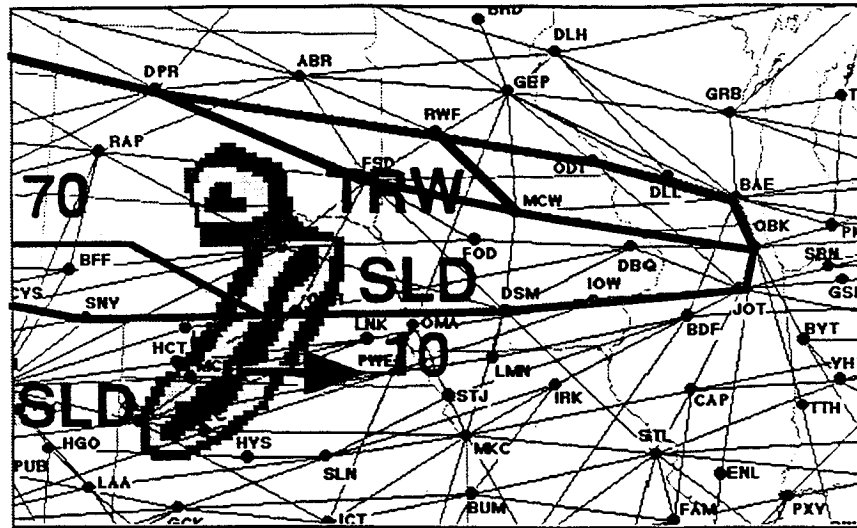
Figure 4. *Fuel-efficient, computer-suggested route versus more conservative northern deviations.*

(a) the subjects observed that the route from DPR to FSD cut close to the forecast storm activity, (b) this observation influenced them to consider the possibility that the forecast might be wrong and that the storm might move further north or east than predicted, and (c) they consequently chose the more conservative path from DPR to RWF.

This observation contrasts with the behaviors of all but two of the subjects in the two route constraints function conditions. These 18 subjects viewed the computer's recommended solution at the national map level (often without even displaying the jet routes) and simply concluded that it looked okay without closely focusing on the choices at DPR. For example, while looking at the national map, Subject C4 stated:

> See if I can get the computer to find a route. [He used the route constraints function with the constraints of no turbulence and no precipitation, and the computer suggested a northerly route.] With a northerly deviation, I can get by with the constraints I placed on it. Now I want to check and make sure. [He observed the destination parameters for es-

timated time and fuel consumption.] This gives me, actually, a shorter flight plan and plenty of fuel at arrival. So I would go ahead and select that route at that point.

Unlike the sketching-only subject described earlier, there is no evidence that Subject C4 considered the uncertainty associated with the storm or that he considered a more conservative northerly deviation.

In short, rather than "explaining" the effects of the automatic display of suggestions with a label such as *overreliance*, it is more informative to conclude the following:

(1) The form of planning required by the sketching-only version forced those pilots to generate plans one leg at a time.
(2) The goal of generating the leg leaving DPR caused the pilots to zoom in on the display shown in Figure 4.
(3) Viewing this display in the context of the goal of selecting the next leg of the flight induced the pilots to consider the uncertainty associated with the storm.

Thus the design of the system interacted with specific characteristics of the scenario to influence the pilots' cognitive processes.

Incomplete predictions. Earlier we described the support of these results from Case 1 for two general predictions that we made. Note that although we made general predictions regarding the possible influence of automatic suggestions on a planner's consideration of the uncertainty associated with a forecast, we did not recognize ahead of time that these predictions applied to the scenario presented in Case 1. (We did not predict that the sketching-only subjects would tend to select a more conservative northerly deviation in this scenario.) As designers, we unintentionally mimicked the behavior of the subjects in the automatic version and did not adequately consider the potential effects of uncertainty in the weather on the subjects in the sketching-only version.

This suggests that it is not sufficient as part of the design process to make general predictions about the effects of alternative designs. To adequately evaluate design alternatives analytically, it is necessary to map out detailed scenario-specific predictions of alternative internal cognitive activities and *sequences* of behaviors that users might exhibit. (This implies that it is also necessary to develop a comprehensive set of scenarios.) Furthermore, these results illustrate the value of empirical evaluations to detect designer oversights.

Initial evaluation of system designs. In Case 1 we can't criticize the computer's suggested route or either of the more conservative northerly routes selected by the pilots in the sketching-only version—all of them are reasonable. We might, however, speculate that in other circumstances, the *cognitive activities* induced by the sketching-only version (if these cognitive processes persist in other scenarios) could lead to more exploration and deeper consideration of the implications of uncertainty in the forecast, leading to the selection of a superior route. (Data relevant to

this hypothesis will be presented in Case 3.)

If this behavior persists in other scenarios, it might be construed as an advantage in the design of the sketching-only system. There was also evidence of behaviors in Case 1, however, in which the sketching-only version put some of the subjects at a disadvantage. In particular, two of these subjects selected a plan that deviated from the original plan at DBS, a second possible deviation point, rather than MYL, the earliest possible deviation point (see Figure 4). This second deviation point is less preferable in terms of fuel consumption.

In abstract terms, then, we again see important effects induced by the system designs. The subjects in the route constraints and sketching and the automatic route constraints conditions let the computer pick a fuel-efficient point for deviation from the original plan. Because of the large solution space, however, the sketching-only subjects were (as predicted) faced with a reasonably difficult task when identifying the best deviation point.

Overview. Case 1 provides clear evidence that the design of the system has strong effects on pilots' performances. More important, it provides insights into the ways in which design features interact with the characteristics of this task (scenario) to influence the user's cognitive processes.

The data from Case 1 indicate that in some ways, use of the computer to produce suggested plans degrades the pilot's role in evaluating plans, whereas in other ways (i.e., finding fuel-efficient solutions) it enhances performance. The following cases provide further data to assess this apparent trade-off between these different design concepts.

Case 2

Case 2 was designed so that there were two initially plausible directions for deviating

90

(north or south of a storm). The scenario consisted of the following:

> It's summer and you are eight minutes into a flight from Oakland to Joliet. You got off the ground at 1600 Zulu. You notice that there is a solid line of convective thunderstorms directly in your path. Decide what you think the aircraft should do. [Figure 5 shows the weather for this case.]

Subject S6: sketching-only system. Prior to the experiment we hypothesized that many of the subjects using the sketching-only version would explore and select a southern deviation in light of the fact that the southern thunderstorm cell appears to be smaller than the northern cell. Because of tailwinds to the north, and headwinds to the south, a northern deviation was clearly preferable in terms of fuel consumption and time of arrival.

After looking at the current and forecast fronts and composite radar (see Figure 5), this subject sketched a southern deviation, compared it with the original route, and checked it for turbulence. Without checking the winds or sketching other alternatives, he selected this deviation.

Subject C3: route constraints and sketching system. Subject C3 looked at the current and forecast composite radar and concluded that he could deviate either to the north or to the south. He decided to let the computer find a deviation based on constraints of light turbulence and light precipitation. The computer suggested the northern deviation shown in Figure 5, and the subject checked it for turbulence. After finding no turbulence along the deviation, he checked it for clearance from the thunderstorms. The subject decided that the distance between the route and the thunderstorms was adequate. He then decided to fly the computer-suggested northern
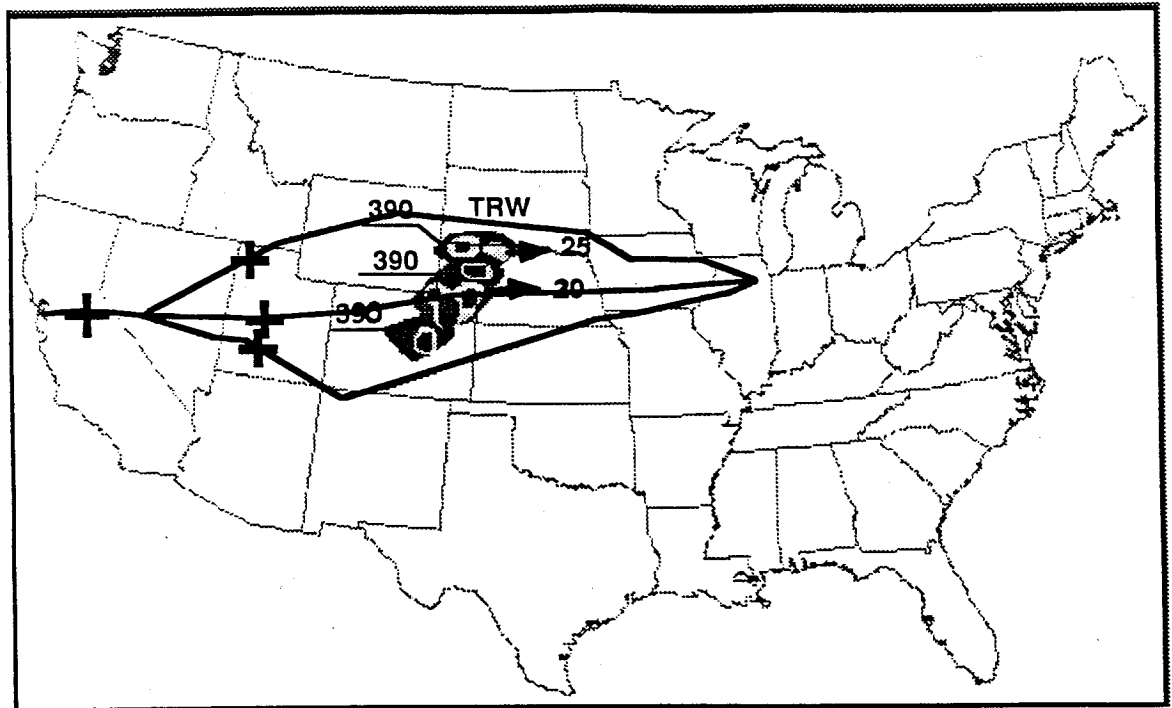

Figure 5. *Routes explored by sample subjects in Case 2 (plotted on a map showing the forecast weather).*

91

route but stated that he would keep an eye on the thunderstorms.

Subject A9: automatic route constraints, route constraints, and sketching system. This subject first looked at the composite radar for the current weather map. He compared the time and fuel consumption for the two routes (the original route and the automatically suggested northern route) and noted their differences. Finally, he gathered some more weather information, including winds, and decided to accept the computer-suggested northern route.

Case 2 Summary Statistics

Whereas 9 of the 10 route constraints and 10 of the 10 automatic subjects decided to deviate north of the original route, 4 of the 10 sketching-only subjects deviated to the south. Based on a chi-square test, this difference was significant at $\alpha < 0.044$. Nevertheless, the sketching-only subjects were not the only ones to explore both northern and southern deviations in detail: 8 of 10 sketching-only, 5 of 10 route constraints, and 4 of 10 automatic subjects explored both northern and southern deviations in detail.

As in Case 1, the information that the subjects looked at was also analyzed on the basis of treatment condition (see Table 1).

Case 2 Discussion

As stated earlier, one of our hypotheses was that because of the large number of possible solutions to explore, subjects in the sketching-only version would be less likely to find the most fuel-efficient route that also avoided the bad weather. This effect was clearly shown in Case 2, in which 40% of the sketching-only subjects selected a southern deviation. The various southern deviations selected used about 3% more fuel and took about 8 min longer than did the northern deviation.

This difficulty in identifying the most fuel-

efficient deviation was in part attributable to a failure to access all of the data in evaluating solutions. Three of the subjects in the sketching-only version failed to look at the map display for winds and consequently did not realize the southern deviation had significant headwinds. Thus, because of the large solution space and the large data space, subjects in the sketching-only version had more difficulty in generating the most efficient route and in evaluating the less satisfactory southern route.

Case 3

Case 3 was designed to present the pilots with a difficult planning problem and to put the various system designs to a demanding test. Unlike the previous cases, the thunderstorms in Case 3 were not localized, and their tops were not all at the same altitude. Like Case 2, there were two likely directions for deviating, but in this case neither was without potential problems. In particular, a deviation that avoided storms at the beginning of the route had to pass through more severe storms later. Finally, flight safety was a bigger concern in this case than in the previous cases.

The following scenario was read to the subjects prior to their working on the case:

> It's summer and you're on a flight from Cheyenne to San Antonio. You got off the ground at 1900 Zulu and are now two minutes into the flight. Decide what you think the aircraft should do.

The original route, the current aircraft position, and the current composite radar are shown in Figure 6. The current radar shows a number of thunderstorm cells with tops ranging from 28 000 to 43 000 ft, but the aircraft's maximum altitude was 33 000 ft. One of the cells directly on the flight path had a top of 43 000 ft. The forecast radar showed that the cells were predicted to move north and slightly east.
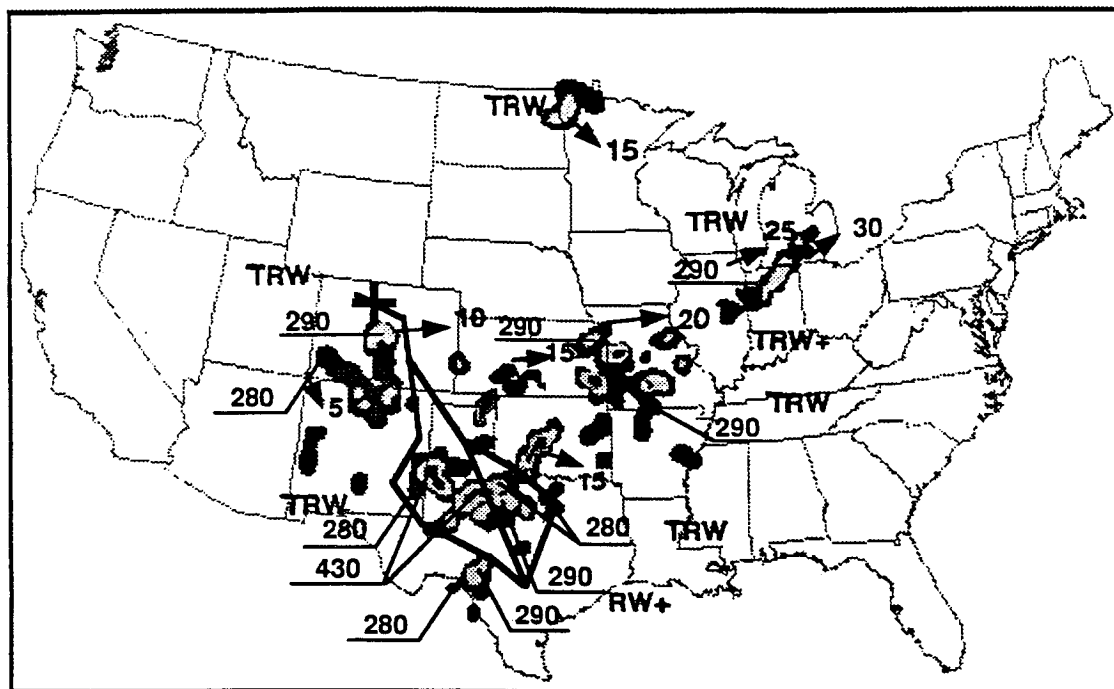
Figure 6. *Original route (the middle route) and suggested alternative routes for Case 3.*

In summary, Case 3 presented subjects with a complex planning problem. The weather was dispersed over a large area and was changing somewhat unpredictably. This scenario required that the pilots anticipate various possible outcomes and plan accordingly. The routes suggested by the computer in the route constraints and sketching and automatic route constraints, route constraints, and sketching conditions are shown in Figure 6. The computer suggested two routes, depending on the constraints placed on it. Constraints of no turbulence and no precipitation caused the computer to suggest the eastern route. Constraints that allowed light turbulence and precipitation caused the computer to suggest the western route. In the automatic route constraints, route constraints, and sketching condition, the route that the computer automatically suggested to the subjects was the eastern route. These subjects had to modify the constraints on the computer or sketch their own route in order to come up with their own western route.

The eastern route passed between two large, severe thunderstorm cells. Summer thunderstorms in Texas are notorious for their volatility, and it was possible that the two cells on either side of the eastern route would grow and build together. Furthermore, the eastern route passed extremely close to a forecast intense cell location.

Prior to the experiment we hypothesized that many of the subjects would have difficulty searching the space of possible solutions and that some of the subjects in the automatic route constraints condition would select the eastern deviation because it was the one initially recommended by the computer, in spite of the fact that it is a questionable choice in both relative and absolute terms. (This case was deliberately selected because the automatic suggestion provided by the computer was poor—poor because the

computer treated the forecasts as reality, rather than reasoning about the uncertainty associated with the forecasts. The weather pattern is, however, realistic. It is based on real weather data provided by the National Center for Atmospheric Research.)

Subject S1: sketching-only system. This subject first indicated that he would have preferred waiting for the weather to clear. Because the aircraft was already en route, however, he considered trying to fly above the weather. He rejected that possibility on seeing the cell tops rising up to 43 000 ft. The subject then spent some time assessing the weather before coming up with two options for dealing with it. First he decided to try going all the way around the back side of the

weather (a far western deviation), but then decided against that option. Deciding to try another western deviation, Subject S1 first tried to deviate from TBE to TCC in order to avoid the cells that lay on the jet route from PUB to TCC (see Figure 7). Realizing that was not possible, the subject tried to avoid the worst of the forecast cells by deviating from PUB to LVS and then back to TCC. After completing the deviation to SAT, the subject compared it with the original route and determined that there was not much fuel remaining.

Subject S1 then looked briefly at a far eastern deviation but instead decided to try a far western deviation around the back of the storm again. After completing the deviation
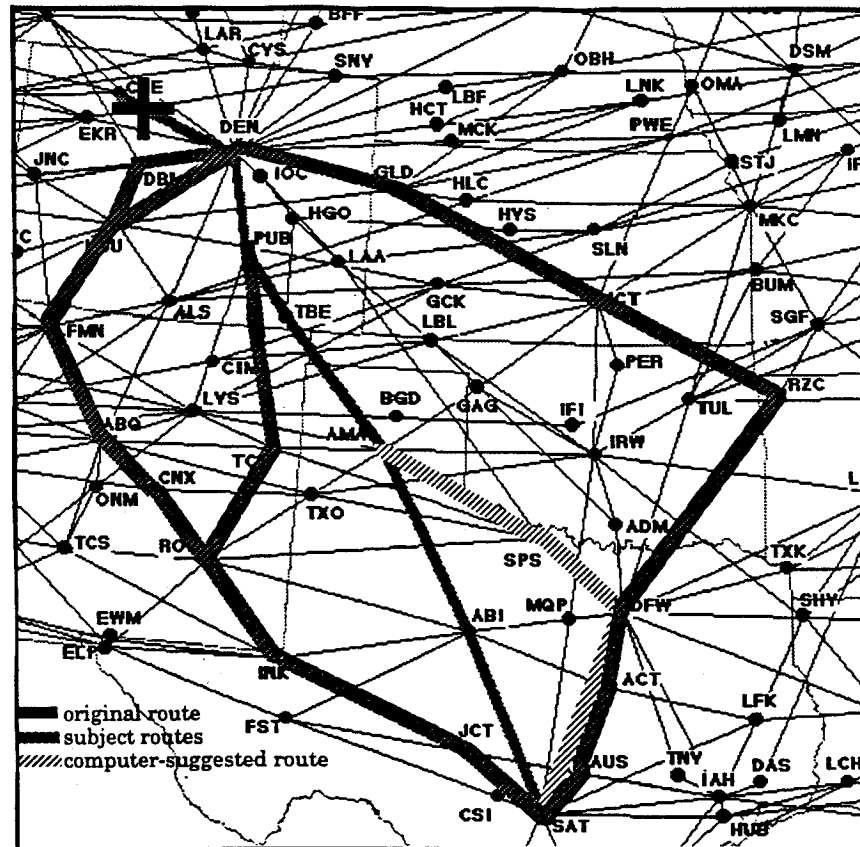

Figure 7. *Routes considered by the sketching-only subjects in Case 3.*

94

and checking it for turbulence, the subject decided that he would continue trying options but that he would start flying a far western deviation. He noticed that this route had increased fuel burn, but the subject also noted that Albuquerque, El Paso, and Dallas were potential alternative destinations. The subject then raised the descent profile from INK to JNC in an effort to avoid the moderate turbulence and to conserve fuel. After comparing the altered profile with the original altitude profile for the deviation, he decided (based on fuel consumption) that he would stick with the original altitude profile.

Finally, Subject S1 sketched another western deviation, but began from AMA rather than PUB. Once again, it appeared as though the subject was trying to avoid the forecast thunderstorm cells south of PUB. Thinking that this route might have saved some fuel, he compared it with the others and noted that the difference was not that large. He then reiterated his choice of a far western route (much farther west than the western route shown in Figure 6).

Subject S6: sketching-only system. Subject S6 spent a fair amount of time assessing the weather before deciding to deviate east from AMA to SPS to DFW (see Figure 7). Like Subject S1, he planned his deviations using forecast weather. In particular, he had zoomed the display around the Denver area when he decided to deviate east. This view clearly showed some moderate thunderstorm cells just south of Pueblo—these likely contributed to his decision to go east. That is, he eliminated possible western deviations from consideration based on a localized criterion or aspect (Tversky, 1972).

Note that this decision was based on forecast conditions, not current conditions; current weather indicated no cells south of Pueblo. This initial decision led the subject to generate and select the eastern route shown in Figure 6, which passed between the two

close, severe thunderstorms near SPS. He did not go back and reconsider his choice of deviation directions to find a more suitable option but announced his intention to fly the eastern deviation. In the debriefing, the subject indicated that the western deviation was clearly preferable (in spite of his choice of the eastern deviation when actually generating his own plan).

Subject C8: route constraints and sketching system. Subject C8 briefly looked at weather information before using the route constraints function with no turbulence and no precipitation. The computer suggested an eastern deviation based on those constraints. After checking the route for turbulence and then further examining the weather, the subject decided to sketch a western deviation beginning with a leg from PUB to TCC (see Figure 7). The subject completed the western deviation (shown in Figure 6) and checked it for turbulence. He then raised the altitude of the leg from INK to JCT (Figure 7) to avoid the moderate turbulence there. Next, the subject tried to find out what was causing the turbulence in the first place. He looked at the destination parameters and indicated that he preferred the western route, then examined the eastern deviation for turbulence and returned to looking at the weather.

The subject next modified the western deviation so that it went from PUB to LVS before returning to TCC (see Figure 7). At this point he spent considerable time examining the two western deviations he had sketched and the weather trends. He finally decided to fly the western deviation that he had sketched first, with the provision that it might have to be modified later depending on how the weather actually developed.

Subject A9: automatic route constraints, route constraints, and sketching system. Subject A9 started by comparing the suggested eastern deviation with the original route (before even looking at weather information).

After noting the differences between the routes in terms of destination parameters and turbulence, he began comparing the routes on the basis of weather. He then sketched a western deviation beginning at AMA and going to ROW (see Figure 7). He checked the turbulence forecast for this western route and rejected that route because it passed through an area where moderate turbulence was predicted up to 29 000 ft for the last third of the flight. He subsequently decided to take the eastern deviation recommended by the computer. In the debriefing he indicated that he would prefer the western deviation over the plan he had selected.

Case 3 Summary Statistics

Table 2 contrasts subjects in terms of whether they selected the computer-suggested eastern route or another route. In addition, as in Case 2, the routes chosen by the sketching-only subjects were much more varied than the ones chosen by the subjects in the other two treatment conditions.

Differences in detailed exploration. Case 3 stands in contrast to the previous two cases, in that the sketching-only subjects did not explore multiple classes of solutions in detail (actually generating specific flight plans) more often than did the subjects in the other two groups. Instead, it was the automatic route constraints, route constraints, and sketching group that explored multiple classes of solutions in detail more often than did the subjects in the other two groups, $\alpha <$

TABLE 2

Final Route Choices for Case 3

	Computer-Suggested Eastern Route	Other
Sketch	1	9
Constraints	3	7
Automatic	4	6

0.022 (5 of 10 in the sketching only, 3 of 10 in the route constraints, and 9 of 10 in the automatic condition explored multiple classes of solutions in detail).

Differences in information search. The number of subjects in each condition who looked at fronts, radar weather, and jet routes is presented in Table 1. There are no clear, statistically significant differences between groups in information searched. The trend, however, seems to be that the route constraints and sketching subjects looked at less information than the sketching-only and automatic subjects.

Case 3 Discussion

Once again, the data indicate that the system design strongly influences the exploration and plan selection processes of the subjects.

Search difficulties. Some of the same challenges in searching the space of possible solutions that occurred in Case 2 recurred in Case 3 for subjects in the sketching-only condition. For example, Subject S10 made six attempts at sketching routes (some were completed, some were aborted) before sketching the route that he finally chose. Similarly, Subject S7 made six attempts at sketching routes before choosing one of them. This difficulty experienced by the sketching-only subjects in generating effective plans was strikingly illustrated by one pilot who developed and chose a deviation all the way east around the entire storm, using up 24% more fuel than the more reasonable western deviation.

Poor search strategies. Subject S6, described in detail earlier, illustrates a fascinating example of how particular strategies can lead to poor solutions. His strategy can be characterized as an elimination-by-aspects approach (Tversky, 1972), in which the aspects are local decisions about which waypoint to go to next.

In particular, he began by saying, "Where

should I go next, from PUB to TCC or from PUB to AMA?" He selected AMA because it was farther from the storm west of TCC. He then considered, "Should I go from AMA to SPS or to ABI or to TCC?" (See Figure 7.) He selected SPS. Because of these localized decisions, he never even considered whether this eastern deviation was to be preferred globally to the western route.

Similarly, several subjects in the automatic version exhibited ineffective strategies. Specifically, they first noted the computer's automatic suggestion of the eastern deviation (see Figure 6). They subsequently generated the western deviation (either by sketching it or by changing the constraints and having the computer generate it). They then viewed the display of predicted turbulence and rejected the western deviation based on the presence of moderate turbulence at some altitudes in the last third of the flight. They did not note any other plausible alternatives, so they accepted the computer's initially suggested (but very poor) eastern route.

Two underlying processes appear to be contributing to this poor performance. First, these subjects are using a single aspect or criterion to reject a plan, rather than evaluating the plan globally in comparison with alternatives. Second, they appeared to accept the computer's poor initial suggestion by default after they rejected the western deviation. In particular, like the pilots in the automatic version in Case 1, they did not show evidence of considering the uncertainty associated with the weather around the eastern deviation.

Disorientation. A final interesting behavior was the failure of some pilots to view the appropriate data when evaluating an alternative. These pilots were looking at the forecast weather while making decisions about the initial segment of the flight. They should have looked at the original weather display to guide decisions that early in the flight. (They appeared to be unaware of which weather display—forecast or current weather—they were looking at.)

Summary. A number of subjects in all three conditions exhibited poor performance in Case 3. Although more subjects appeared to be biased toward a poor solution when it was suggested by the computer, this bias cannot be explained simply by overreliance in the sense of blindly accepting the computer's recommendations. These subjects showed clear evidence of generating and evaluating alternatives. Thus much deeper explanations had to be developed to account for their acceptance of the computer's poor suggestion. In addition, forcing the pilots to be more involved by making them sketch their own solutions resulted in the selection of fewer poor plans. Nevertheless, because of the use of an elimination-by-aspects strategy by one subject, he generated and selected the poor eastern deviation without any suggestions from the computer.

Case 4

Case 4 presented subjects with a situation in which the shortest and most fuel-efficient deviation, north, required the pilots to violate one of their standard heuristics (fly upwind of thunderstorms). The storm in this case could also be topped though that would have put the aircraft in turbulence above the storm. Furthermore, there was some risk of the storm growing quickly. As in the previous two cases, there were two likely directions for deviating; in this case those directions were north and south of the storm.

The following scenario was read to the subjects prior to their working on the case:

> You are on a flight from Albuquerque to New Orleans. You got off the ground at 1400 Zulu. You are now 19 minutes into the flight and have noticed a thunderstorm cell outside of Dallas. Decide what you think the aircraft should do.

The original route, the current aircraft position, and the current composite radar are shown in Figure 8 along with the likely deviations north and south of the storm. The forecast weather showed the storm moving slowly to the northeast.

Subject S1: sketching-only system. Subject S1 began by looking at weather information and may have considered flying over the top of the weather; he wondered aloud how high the cell went and noted that it went up to 28 000 ft and that the aircraft was planned to fly at 33 000 ft. On noticing moderate turbulence at 33 000 ft, however, the subject decided to try a southern route. After sketching a southern deviation, he checked it for turbulence and compared it with the original route. The subject then reviewed the weather and the original route and sketched a route to the north. After checking the route for turbulence, he decided that either route would work. Because the storm was isolated, he decided to take the route that consumed the least fuel, which was the northern route (even though he noted that the storm was moving north).

Subject C8: route constraints and sketching

system. Subject C8 started by checking the weather and then decided to use the route constraints function to find a new route based on constraints of no turbulence and no precipitation. The computer suggested a northern deviation, which the subject checked for turbulence and compared with the original route. Subject C8 then decided to sketch a southern deviation to see if it was any better. In comparing the northern and southern routes, he noted the trade-off between the two: The northern route took less time, but the storm was slowly moving in that direction. He rechecked the weather and the turbulence on the southern route and stated that it did not matter which one he chose.

He was continuing to look at the weather and the southern route when he noticed the possibility of flying above the weather. This prompted him to relax his constraints to light turbulence and light precipitation and try the route constraints function again. The computer again suggested the northern route. The subject then went through a process of reasoning about the uncertainty of the forecast, the position of the aircraft relative to the weather, and the costs of avoiding all of the
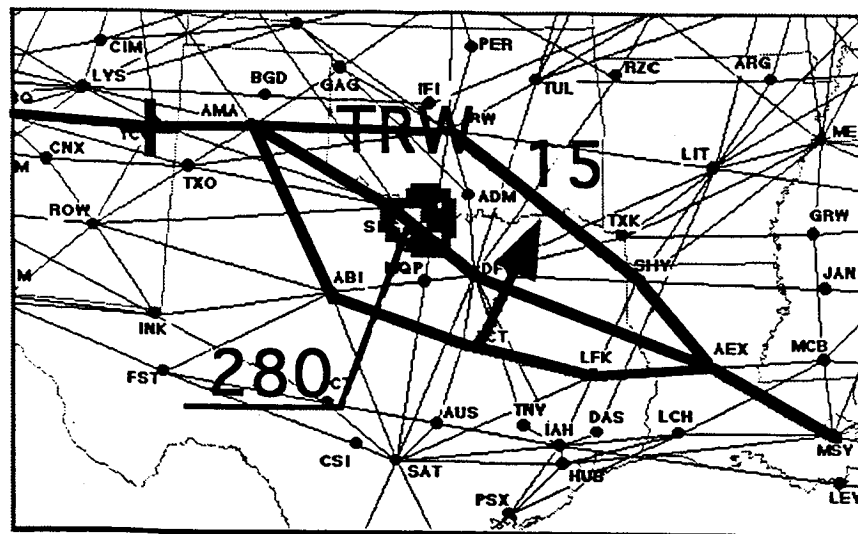

Figure 8. *Current composite radar for Case 4.*

uncertainty. This resulted in his choosing the southern deviation so that he would be assured of avoiding the weather.

Subject A7: automatic route constraints, route constraints, and sketching system. Subject A7 first compared the destination parameters of the original route and the computer-suggested northern deviation. He then checked the deviation for turbulence and continued investigating the weather. Subject A7 decided to try a southern deviation and compared the destination parameters of that route with those of the other routes. After gathering more weather information, this pilot diverted to the south.

Case 4 Summary Statistics

The following data are for all 30 subjects.

Differences in final routes. For the most part, there were three reasonably likely route alternatives: north of the original route, south of the original route, and the original route (at a higher altitude). Subjects were grouped on the basis of choosing a route similar to the computer-suggested northern deviation, a southern deviation, or the original route. This analysis is presented in Table 3. In terms of fuel consumption and time of arrival, the northern deviation is slightly better than the southern one: A southern deviation consumes 368 lb (167 kg) more fuel (about 2% of the fuel consumed by the northern deviation) and takes an additional 3 min. However, isolated thunderstorms in Texas are sometimes called *super cells* because of their volatility and unpredictability. A third of the

TABLE 3

Final Route Choices for Case 4

	North	South	Original
Sketch	5	4	1
Constraints	5	4	1
Automatic	7	2	1

subjects believed the trade-off in time and fuel was worth the added security of a southern deviation (because the storm was moving northeast).

Differences in detailed exploration. Case 4 is similar to Case 3 in that the sketching-only subjects show a trend not to explore multiple classes of solutions in detail as often as do subjects in the other two groups. This trend is only marginally significant, $\alpha < 0.142$ (6 of 10 sketching-only, 5 of 10 route constraints, and 9 of 10 automatic subjects explored multiple classes of solutions in detail).

Differences in information search. The number of subjects in each condition who looked at fronts, radar weather winds, and jet routes is presented in Table 1. There were no clear differences in the groups.

Case 4 Discussion

Unlike the previous cases, the data do not provide strong evidence that the system design strongly influenced the exploration and plan selection processes of the subjects, though there is a trend showing that the automatic route constraints subjects chose the computer's suggested route more often. This may be attributable to the much smaller solution space present in Case 4.

Individual differences. Although the data again suggest a possible (nonsignificant) biasing effect attributable to the computer's automatic suggestion, the primary result of interest is the evidence that pilots differ in their evaluations of alternatives. Some clearly preferred deviating north to save time and fuel. Others preferred the more conservative southern deviation to decrease the likelihood of encountering the storm. (The available data are not informative regarding the causes of such differences. Preference differences could result from different mental models of the weather or air traffic, differences in utility functions, or differences in training or flying experiences, among others.)

99

CONCLUSION

Our results show that, at least for the foreseeable future, it will be infeasible to fully automate tasks such as en-route flight planning given the current state of technology. Feasible methods for adequately dealing with reasoning about such complex, uncertain events and for considering the trade-offs among goals such as safety, cost, and passenger comfort simply do not exist at present.

However, current and developing technologies seem to offer interesting opportunities for enhancing flight planning activities, such as the following:

(1) Designing better interfaces to provide more perspicuous displays of such data and information and to incorporate graphical interfaces that allow direct manipulation of routes to explore alternatives
(2) Providing access to more complete and accurate information on weather, air traffic, and airport conditions in a timely fashion
(3) Using optimization and expert systems technologies to assist users in generating and evaluating alternative plans and to provide intelligent alerting functions
(4) Using the computer to enhance communication and cooperation among the various people concerned with flight planning (dispatchers, flight crews, air traffic control, etc.)

In our research we have encountered evidence to support all four points. The present study is particularly important in raising issues regarding the first and third areas, however. The results clearly demonstrate the critical need to address the question: How can advanced technologies be applied to develop cooperative planning systems that effectively support the activities of users?

In spite of the emphasis in this paper on errors induced by the Flight-Planning Test Bed, overall, the design features of all three versions supported successful efforts. In Cases 1, 2, and 4, all of the plans selected using all three versions of the system were acceptable, though some were less efficient in terms of fuel consumption and flight time.

The overall efficacy of the design of the Flight-Planning Test Bed as a cooperative system was further supported by the reactions of the pilots:

"I think it's great. It gives you another piece of information to consider. It's like delegating responsibility."

"It would be great if you could sit down with your dispatcher and do this sort of thing before a flight."

"I like it. Being able to zoom in on the route and look at the weather and the projections is nice. It's pretty easy to use. It's pretty straightforward. It's got everything you need."

"I wish we had something like this now, especially in operations. You'd have to kill guys to get them off of it."

"I'm pretty impressed by this. If you could get the lunch menu on here too, you'd have it made!"

Nevertheless, as Case 3 most dramatically demonstrated, certain design features can induce unacceptable performances.

There will no doubt be a strong temptation to let technology drive the development of future flight-planning systems because the potential value of the available computer and telecommunications technologies seems so apparent. This study, however, provides strong evidence that the design of the computer support system can clearly influence the exploration and evaluation of alternative system plans by users. The data demonstrated that even when alternative designs provide access to the same data, some designs can exert powerful and undesirable effects on the problem-solving processes of the user and on the final product of these processes (the selected flight amendment).

The following summarizes the various undesirable effects observed in this study and discusses recommendations for system designs and future research.

Large Data Spaces

In the near future it will be possible to provide flight planners with access to a rich set of data relevant to the planning process. As this study illustrated, however, more is not necessarily better. Even with the limited sources of data available to users of the Flight-Planning Test Bed, we saw evidence of disorientation and failure to attend to important data. Such effects are likely to increase as we provide access to even more data displays.

Some pilots, for instance, failed to recognize that they were looking only at the forecast weather when planning early segments of the flight (in situations in which current weather displays were clearly relevant). The result for one subject (in the sketching-only version) in Case 3 was to completely overlook the best solution and to accept a poor flight plan.

In addition to such disorientation, some pilots failed even to look at important data such as the winds. This was a major contributing factor leading to the selection of the less desirable southern deviation in Case 2 by several sketching-only subjects.

Several design principles are suggested by such data:

(1) When designing the computer system, select the data to display judiciously. Providing access to more kinds of data, even though in principle they may be useful, does not ensure that they will be used effectively at the right time.
(2) Develop good representations to make the implications of important data and relationships salient to the user. (One interesting example of a problem was discovered with current displays of wind information: A number of pilots did not know how to interpret wind charts regarding the strength or direction of the winds.)
(3) Consider designing integrated data displays to communicate information pertinent to a common goal for which those types of data are pertinent. The Flight-Planning Test Bed demonstrates the integration of weather data

with displays of alternative flight paths. Displays that integrate data on precipitation and turbulence at different altitudes would similarly be useful (but not trivial to design).
(4) Provide clear feedback about the state of the display (such as whether the displayed data represent current or forecast weather). It is not enough to present such data on the state of the display. It must be highly salient.
(5) Consider incorporating intelligent alerting functions to ensure that critical data (or the implications of these data) are not overlooked.

Large Solution Spaces

Because of the large number of possible flight paths, the subjects in the sketching-only version sometimes had difficulty finding a good alternative. In circumstances in which time is critical, such difficulties could also use up valuable time and attention.

This problem suggests the potential value of tools (based on optimization or expert system technologies) to help search for good solutions. Indeed, without such tools, the subjects frequently found solutions that used up significantly more flight time and fuel and were no better in terms of other criteria. One subject, for instance, selected a plan in Case 3 that used up 24% more fuel.

A counterargument to utilizing such technologies is that they are brittle. They may be good for routine situations that the designer has anticipated, but they also fail in unanticipated situations. Such a line of argument continues by suggesting that we keep users "in the loop" by making them do more of the work and by suggesting that, because they must therefore stay involved, they will notice and deal with unusual situations. In short, this argument suggests that although people won't always find the best solution, by keeping them involved bad solutions are avoided. Clearly, the extreme form of this argument is a "straw man." People make errors, too. Consequently, the trade-offs must be weighed between the potential errors made by the designer (including those for situations that are

not known, because otherwise the designer could design for them!) and errors made by users. Also, designs must assume that both designers and users are fallible.

Case 3 illustrated the fact that keeping the person in the loop does not ensure that poor solutions will be avoided: One of the sketching-only subjects generated and selected the poor eastern route on his own. Thus a principle such as "avoid excessive automation in order to keep the person involved in the task" is too simplistic. Keeping the person involved does not ensure more exploration, nor does it ensure that solutions will be chosen that are at least satisfactory. Instead, how specific types of designs will interact with users' cognitive processes in specific types of scenarios to produce undesirable behaviors must be considered. (The discussions of results for Cases 1 and 3 provide pertinent illustrations.)

In terms of this application area, what is needed is a design that lets the computer use its power to help search the solution space while keeping the person involved and while protecting against errors the person may make. The first two problems might be addressed by developing sophisticated perceptual displays that make alternatives easier to generate and evaluate. (One possibility would be a display that allowed the user to prune undesirable jet route segments by setting constraints, for example, telling the computer to hide or dim all jet routes on the map that pass through more than light turbulence.) Another solution might be using optimization or expert systems technologies to let the computer generate alternatives (which is what the Flight-Planning Test Bed does using the route constraints tool) but improving the design by having the computer generate the best alternative(s) for each class of solutions and then letting the user evaluate these alternatives. Thus the computer might display the "best" deviations both north and south of a storm for comparison by the user. We specu-

late that both of these potential solutions would keep users involved because they would have to look at the data to make choices among alternatives.

A solution to the third problem (protecting against the user's errors) is more complicated, though, because the nature of the errors the person might make must first be predicted.

Overreliance

As pointed out earlier, system users sometimes develop poor plans even when they are kept in the loop. (Case 3 illustrated this behavior.) However, because of the limitations of the technology used in the Flight-Planning Test Bed, the route constraints function also produced a poor suggestion in Case 3. (This resulted from the fact that the Flight-Planning Test Bed does not reason about the uncertainty associated with forecasts.) Our study illustrated that even though subjects in the automatic suggestion version used the available "manual" functions to explore alternatives to the computer's suggestion, 40% still wound up accepting this poor plan. Based on this result, it is worth emphasizing that the effects of providing automatic suggestions by the computer can be quite pronounced. Subjects in Cases 1 and 3 who were presented with the computer's suggestion clearly reasoned less (or not at all) about the uncertainty associated with the forecast, leading them to accept a poor flight plan in Case 3. Also, labels such as *overreliance* are too shallow as descriptions or explanations of user performance to adequately guide design decisions. The design of the computer influences users in a complex, scenario-specific fashion. (The discussions of the results for Cases 1 and 3 present illustrations of such effects.) Thus, to evaluate proposed support tools, scenario-sensitive cognitive models need to be considered.

This failure by subjects to reason about un-

certainty when viewing the computer's suggestion might be alleviated by either of the two solutions outlined earlier. Although the sketching-only version induced subjects to look at critical data and question which path is better if the forecast is wrong, requiring them to choose from among several alternatives suggested by the computer might induce them to look at the critical data and ask the same question.

A further form of protection against such failures to consider uncertainty would be the incorporation of an intelligent alerting function that either warned the person when a route might be too close to a developing problem or inhibited the display of a suggested route by the computer if it appeared to be too close to a developing problem (thus making the computer conservative in suggesting alternatives). One caution is in order regarding these potential solutions, however. Subjects may fixate on the set of alternative solutions suggested by the computer and consequently fail to note an even better solution that the computer missed, or fail to note that the computer has suggested a poor solution.

Maladaptive Strategies

The literature on human problem solving provides numerous examples of how people apply simplifying heuristics in order to reduce the complexity of a decision (Elstein, Shulman, and Sprafka, 1978; Fraser, Smith, and Smith, 1992). One such strategy is to eliminate an alternative based on a single criterion, rather than evaluate the alternative more globally (in terms of trade-offs among all of the relevant criteria). In Case 3 this type of strategy was exhibited by subjects' using all three versions of the system. The result was the selection of a poor plan by 10% of the subjects in the sketching-only version, 30% of the subjects in the sketching and route constraints version, and 40% of the subjects in the automatic version.

Having the computer indicate several possible solutions might help to encourage a more global evaluation. In addition, it might be helpful to use animation to create displays to help the user view the data over the entire flight and to include redundancy in the evaluation of plans (e.g., letting the flight crew look at displays of paths proposed by a dispatcher, or vice versa).

Supporting Individual Differences

Results such as those in Case 4 provide strong evidence for the need to give people the option to explore alternatives on their own. Because people differ in terms of their preferences and mental models of a situation and because there is no objective way to say who is right for each situation, designers need to give people the tools necessary to allow them to create their own alternatives and to play "what if" games, even if the computer provides some suggestions.

Final Note

This study demonstrates that the design of an effective cooperative system for a complex task such as flight planning is a significant challenge. It requires careful consideration of how system design features and specific displays influence the cognitive processes of users in specific types of scenarios. In general, the results highlight considerations for designers of cooperative problem-solving systems, particularly those that support dispatchers and flight crews.

Are such considerations worth the effort? The ability of a system design to induce 40% of the pilots studied to select a poor flight plan suggests that there is indeed a real need to explore these issues further and to take them seriously when implementing commercial systems.

ACKNOWLEDGMENTS

REFERENCES

Cohen, M. S., Leddo, J. H., and Tolcott, M. A. (1989). *Personalized and prescriptive aids for commercial air flight replanning* (Report 89-2). Reston, VA: Decision Science Consortium, Inc.

Coombs, M., and Alty, J. (1987). Expert systems: An alternative paradigm. *International Journal of Man-Machine Studies, 20,* 21–43.

Elstein, A., Shulman, L., and Sprafka, S. (1978). *Medical problem solving: An analysis and clinical reasoning.* Cambridge, MA: Harvard University Press.

Fraser, J. M., Smith, P. J., and Smith, J. W., Jr. (1992). A catalog of errors. *International Journal of Man-Machine Studies, 37,* 265–307.

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

Hoc, J. M. (1988). *Cognitive psychology of planning.* London: Academic.

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

Lehner, P. E., and Zirk, D. A. (1987). Cognitive factors in user-expert-system interaction. *Human Factors, 29,* 97–109.

Miller, G. A., Galanter, E., and Pribram, K. H. (1960). *Plans and the structure of behavior.* New York: Holt.

Robertson, S., Zachery, W., and Black, J. (Eds.). (1990). *Cognition, computing and cooperation.* Norwood, NJ: Ablex.

Rudolph, F. M., Homoki, D. A., and Sexton, G. A. (1990). *"Diverter" decision aiding for in-flight diversions* (Contractor Report 182070). Hampton, VA: NASA.

Sacerdoti, E. D. (1974). Planning in a hierarchy of abstraction spaces. *Artificial Intelligence, 5,* 115–135.

Schank, R., and Abelson, R. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Erlbaum.

Shute, S., and Smith, P. J. (1993). Knowledge-based search tactics. *Information Processing and Management, 29,* 29–45.

Smith, P. J., McCoy, E., Layton, C., and Bihari, T. (1992). *Design concepts for the development of cooperative problem-solving systems* (OSU Tech. Report CSEL-1992-07). Columbus, OH: Ohio State University, Department of Industrial and Systems Engineering.

Sorensen, J. A., Waters, M. H., and Patmore, L. C. (1983). *Computer programs for generation and evaluation of near-optimum vertical flight profiles* (Contractor Report 3688). Hampton, VA: NASA.

Stefik, M. (1981). Planning with constraints: MOLGEN part 1. *Artificial Intelligence, 16,* 111–140.

Suchman, L. A. (1987). *Plans and situated actions: The problem of human machine communication.* New York: Cambridge University Press.

Thierauf, R. J. (1988). *User-oriented decision support systems: Accent on problem finding.* Englewood Cliffs, NJ: Prentice Hall.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79,* 281–299.

Wilensky, R. (1983). *Planning and understanding: A computational approach to human reasoning.* Reading, MA: Addison-Wesley.