In R. Jensen (Ed.), Proceedings of the 8th International Symposium on Aviation Psychology, April 1995, Columbus, OH, 777-872.

EXPERT PILOTS' PERCEPTIONS OF PROBLEM SITUATIONS [1]

Ute Fischer — San Jose State University
Judith Orasanu — NASA-Ames Research Center
Mike Wich — San Jose State University

How do expert pilots interpret flight-related problem situations and what aspects do they find important for making decisions? These questions are not as easy to answer as they at first might appear. Aviation decision situations are often complex and a number of factors may influence decision making. Candidate variables are - to name just a few - phase of flight, aircraft type, problem type, available time and resources, or risk involved. Orasanu (1994) presents a model of aviation decision making that emphasizes the role of situation assessment: Deciding on a course of action is contingent upon an adequate understanding of the nature of the problem and of the response options afforded by the problem situation. Risk and time pressure are situational variables that further constrain the decision process. One goal of our present work was to validate this process model from the pilot's perspective.

Our second goal was to determine the role of pilot experience on the selection of decision relevant features. Numerous studies on expertise have shown that experience changes people's understanding of problems (e.g., Chase & Simon, 1973; Chi, Feltovich & Glaser, 1981). Similarly we expected that more experienced pilots approach situations differently and focus on different aspects than less experienced pilots. Experience in an aviation context, moreover, does not only mean years of pilot practice but also has a more specific sense: experience in a particular crew position. Captains, first officers and flight engineers fulfill different roles and have different responsibilities during both normal and abnormal situations. How pilots view problem situations may thus be a function of their role-specific experience rather than be determined by years of experience.

In studies 1 and 2, we investigated what features pilots select during a hypothetical decision making exercise and whether experience influences this process. Study 3 was designed to explore further the role of experience in this context.

STUDY 1

METHOD

Material

Ninety-seven incident reports on decision making were retrieved from the Aviation Safety Reporting System (ASRS) data base. A stratified random sample of 22 scenarios was selected to reflect phases of flight in the report set. The following distribution was obtained: There were 3 scenarios in which the incident occurred during pre-flight. Two scenarios involved incidents during taxiing. Four scenarios had the incident during take-

---

[1] Part of this work was presented at the Annual Meeting of the Society for Judgment and Decision Making in St. Louis, November 1994. The research described in this paper was made possible through the support of NASA-Code UL and the FAA-ARD-210. We wish to thank the participating airlines, especially the pilots, who generously shared their time as subjects. We are also indebted to Don Bryant who helped us as subject matter expert throughout the project, to Jerry Jones, Corwin Logsdon, Mietek Steglinski, and Barbara Sweet who volunteered as preliminary subjects, and to Kathy Mosier who commented on an earlier version of this paper.

off, another 4 during climb and 4 more during cruise. Incidents during descent were described in 2 scenarios; during approach in 3 scenarios.

The scenarios were shortened to suit our purpose. All the information that the reporter had provided regarding flight context and precipitating event was included. The actual decision, however, was not mentioned. Subjects, for instance, saw the following description: *A wide bodied aircraft weighing 43,000 lbs experiences a #3 engine compressor stall at 138 kts on takeoff; V1 is 142 kts.* Each description was printed on an index card.

Procedure

Subjects received the stack of 22 cards and were asked to infer the appropriate decision for each incident and then to group together those scenarios that involved the same kind of major decision. Subjects were told that they could create between 2 and 21 piles, and that they could do multiple sorts as long as they used different bases in subsequent trials. Each sort, however, should reflect similarity with respect to decision making. Subjects were tested individually and were allowed to work at their own pace. Upon completion, the sortings were recorded and subjects were asked to describe the common element of those scenarios that they had placed together in a pile. Subjects who provided more than one sort were asked to choose the one they preferred.

Subjects

Twenty-eight professional pilots from a major US airline volunteered as subjects. There were 10 captains, 10 first officers and 8 flight engineers. The captains had on average 27 years of experience in civil transport, first officers 15 years, and flight engineers 5 years.

Analyses

Multidimensional Scaling and Hierarchical Clustering analyses were conducted on subjects' sortings. Only data from single or preferred sorts were used. Separate analyses were performed on the groupings of captains, first officers, and flight engineers. Only findings from the Multidimensional Scaling analysis will be reported in this paper. The interpretation of dimensions was aided by the verbal descriptions that the subjects had provided.

RESULTS AND DISCUSSION

Captains and flight engineers sorted the 22 scenarios on average into 5 piles. First officers created on average 6 piles. Multidimensional Scaling and Hierarchical Clustering analyses converged on a two-dimensional solution for all three groups. However, the interpretation of the dimensions suggests differences between the three crew positions.

Captains

The situational aspects discerned for captains were “Potential Risk” and “Time Pressure,” in descending order of salience. Interestingly, risk was categorized along a time axis with the present as the point of reference. Captains distinguished between situations that involved no risk, and situations that could be risky at some future time. Between these two extreme poles fell situations in which risk was immediate. Examples of each category are:

> ***No risk**: An aircraft fueler advises the crew of a medium large transport that he can’t get the numbers to balance between the fuel off loaded from the truck and the fuel on the aircraft. He works on it for an hour and gets nowhere.*

> ***Immediate risk:*** *A wide bodied aircraft weighing 43,000 lbs experiences a #3 engine compressor stall at 138 kts on takeoff; V1 is 142 kts.*
>
> ***Potential risk:*** *According to its logbook, a medium large transport had not been flown for 10 days; during that time maintenance was done and a new interior put in. On an approach at BUR the right main gear "unsafe" red light comes on and stays on during gear extension with no green down and locked light. At 8-10 miles out, flight level 3000-4000', 180 kts, IAS with 5 degree flaps, the crew recycle the gear but still get the same result.*

"Time Pressure" was the second situational aspect crucial to captains' decision making. It was also the only dimension shared by captains, first officers and flight engineers. Ratings ranged from low to high.

> ***Low time pressure:*** *After all the passengers board a medium large transport at CHA, a ramp agent reports to the flight crew that the forward cargo door will not lock closed. CHA is a non-maintenance base. A mechanic contracted from the airfield, however, is able to get the cargo door latch mechanism to function marginally--i.e., it can only be closed/opened once and the "Fwd Cargo Door" light stays on.*
>
> ***High time pressure:*** *A medium large transport is cleared for takeoff on runway 13C. Power is advanced and takeoff roll commences with the crew setting takeoff power. At approximately 100 kts, a crew member notices a light twin engine taxing at a fairly high speed toward 13C. At the same time, he hears the tower controller making repeated unsuccessful attempts to contact the small plane.*

First Officers and Flight Engineers

First officers were found to attend primarily to "Time Pressure" and then to a variable we may call either "Response Determinacy" or "Situational Complexity." Since the "Time Pressure" dimension corresponded to the captain dimension, it need not be addressed any further. The variable "Response Determinacy/Situational Complexity" could vary from simple to complex. At one extreme were incidents that were clear-cut. Decision making in these instances was considered to be simple: You know what to do, because there is a rule, a procedure, or a policy in place that tells you what to do.

> ***Simple Problems:*** *(1) A large transport is preparing to depart DCA for a flight to MSP. The flight is running late and the DCA curfew is nearing. After push back, it is found that there are 185 passengers aboard, but only 184 seats.*
>
> *(2) A medium large transport is cleared by DFW approach control to descend to 3000' and intercept the 13R ILS. Clearance is also given to contact the tower. On the approach when contact cannot be established via the frequency set on the VHF Com, the crew dial in the frequency for the west tower on the approach plate. Several attempts are made to contact the tower but due to the extreme congestion on the frequency, no landing clearance is obtained.*

Scenarios that were rated more complex formed the other extreme. These incidents were ambiguous either with respect to the nature of the problem or concerning the course of action. Often the crew had to take into account a number of constraints.

> ***Complex Problems:*** *(1) A medium large transport is in cruise at flight level 310, westbound over the North Atlantic, when the crew discover that the crew oxygen bottle is reading 800 psi. The bottle had been changed in Frankfurt due to an incoming crew write-up and had been definitely noted as full. The crew shut off the crew regulators to see if a leak point can be determined; none is found but the supply decreases to 750 psi.*
> *(2) A medium large transport climbing through FL 280 at .82 mach experiences a high frequency flutter through the air frame. Upon visual inspection, it is discovered that the right outboard aileron balance tab is fluttering 1-2" up and down. The flutter stops some time during*

> *deceleration to .74 mach and is stable on second inspection. A company mechanic on board indicates there are no visual signs of delamination or structural failure.*

Flight engineers differed from first officers in the relative weight of the dimensions. For flight engineers we found “Response Determinacy/Situational Complexity” to be the primary dimension followed by “Time Pressure.”

STUDY 2

METHOD

Material and Procedure

Study 2 was conducted to confirm our interpretations of the Multidimensional Scaling analyses in study 1. Instead of using a free sorting task, we employed this time a directed sorting task. Subjects received the same 22 scenarios as before, but unlike study 1, we provided the basis for sorting. Subjects were asked to judge how much risk there was, and how much time pressure; additionally, they were asked to judge how clearly defined the problem was and how straightforward it was to decide on a course of action. If our previous inferences were correct, then subjects who rated the scenarios, for instance, with respect to time pressure, should align them similarly to the “Time Pressure” dimension hypothesized in study 1. All subjects sorted according to risk, time pressure, problem definition and response determinacy. For each sort they were asked to align the piles from the one having least of a particular quality to the one having most of it. The bases for sorting were presented in a randomized order.

Subjects

Nine captains, 10 first officers and eight flight engineers of the same airline as in study 1 volunteered as participants. Captains had on average 24 years of experience in civil transport; first officers 10 years, and flight engineers 6 years.

Analyses

Subjects' sorts were translated into values on a scale ranging from -1 to +1. For instance, the following mapping was implemented when a subject created 5 piles: Scenarios that were judged to be lowest concerning a given criterion received the score -1; scenarios in the second pile were scored as -.5, those in the third and fourth piles were scored as 0 and +.5, respectively, and those in the fifth pile as +1. Mean rating scores were computed for each scenario on risk, time pressure, problem definition and response determinacy. Separate mean scores were calculated for captains, first officers, and flight engineers and used as independent variables in multiple regression analyses to predict scenario values along the dimensions obtained for each group in study 1.

RESULTS AND DISCUSSION

Captains created on average 5 piles when sorting for risk and time pressure, and 4 piles when sorting for problem definition and response determinacy. First officers had on average 5 piles for risk, time pressure and response determinacy, and 4 piles for problem definition. Flight engineers made on average 5 piles for risk and time pressure, 3 piles for problem definition and 4 piles for response determinacy.

Captains. The dimension labels inferred for captains in study 1 were confirmed in several Multiple Regression analyses which used captains' mean ratings obtained in study 2 as predictor variables for the dimensional values. Ratings of time pressure were the best predictor for the “Time Pressure” dimension accounting for 70% of the variance in the values along this dimension ($R^2 = .70$; $F(1,20) = 45.91$, $p < .0001$). “Response Determinacy” was also found to be a component of this dimension. Including ratings of response determinacy as predictor variable, improved the fit of the regression model by 6% ($R^2 = .76$; $F(2,19) = 30.09$, $p < .0001$).

For the captains' “Risk” dimension, a nonlinear regression model consisting of a linear component of risk and a quadric component of time pressure was most suited ($R^2 = .60$; $F(2,19) = 14.37$; $p < .0001$). This pattern of findings supports our previous interpretation of the dimension as “Potential Risk:” Situations that involve risk but time to deal with the problem, are potentially risky and are thus located at the extreme of the “Potential Risk” dimension. Risky situations that require immediate action, have correspondingly lower values on this dimension.

First Officers. Ratings obtained from first officers in study 2 supported only the “Time Pressure” dimension hypothesized in study 1. Seventy-seven percent of the variance in the scores along this dimension were accounted for by ratings of time pressure ($R^2 = .77$; $F(1,20) = 67.25$, $p < .0001$). Ratings for risk accounted for an additional 7% of the remaining variance indicating that “Risk” was part of this dimension ($R^2 = .84$; $F(2,19) = 50.43$, $p < .0001$). Values for the dimension labeled “Response Determinacy/Situational Complexity” could not be predicted by the ratings collected in study 2.

Flight Engineers. Both dimensions hypothesized for the flight engineers in study 1 were confirmed: Ratings of response determinacy were the single best predictor for the flight engineers' first dimension, which should consequently be labeled “Response Determinacy” rather than “Response Determinacy/Situational Complexity” ($R^2 = .20$; $F(1,20) = 4.88$, $p < .05$). The “Time Pressure” dimension was only predicted by ratings of time pressure ($R^2 = .69$; $F(1,20) = 43.81$, $p < .0001$).

Studies 1 and 2 show that aviation decision making involves the following components: Pilots assess how much risk there is in a given situation, how much time they have left for making a decision and whether there is only one, clear-cut answer to a problem, or whether the situation leaves options. We also found that not all pilots used the same set of situational features in making aviation decisions. The captains in study 1 were primarily sensitive to risk and time pressure. Flight engineers and to a lesser extent first officers categorized situations in terms of available response options.

These findings leave us with two questions: (1) Did the captains, first officers and flight engineers in study 1 pick up different situational features because pilots in different positions bring also different kinds of knowledge to decision making? Or, do all pilots know the same things and only different aspects of decision situations are salient to them? (2) Do the differences in feature selection that we observed for captains, first officers and flight engineers, reflect biases specific to crew roles, or are they the result of differences in years of aviation experience? It seems likely that the participants in study 1 focused on those situational aspects that were consistent with the responsibilities associated with their crew position. Alternatively, the differences in feature selection may simply reflect differences in years of experience: The captains in study 1 were much more senior than the first officers who, in turn, had more aviation experience than the flight engineers. We will take up this issue in study 3. First we want to turn to question 1.

Recall that all pilots in study 2 judged the decision scenarios with respect to risk, time pressure, problem definition, and response determinacy. If knowledge differences existed between these groups, captains, first officers, and flight engineers should differ in their ratings of the scenarios.

We computed the correlation coefficient between captains', first officers', and flight engineers' mean rating scores for each scenario and situational variable. As can be seen in Table 1, the groups agreed strongly in their ratings of the scenarios along all four variables. These results suggest that the captains, first officers and flight engineers in study 1 focused on those situational features that were salient to them. Knowledge differences between the groups does not seem to be a plausible explanation.

TABLE 1
CORRELATION MATRIX OF RATINGS

RISK

	CA	F/O	F/E
CA		.85***	.83***
F/O			.96***

TIME PRESSURE

	CA	F/O	F/E
CA		.98***	.93***
F/O			.94***

PROBLEM DEFINITION

	CA	F/O	F/E
CA		.65**	.59*
F/O			.77***

RESPONSE DETERMINACY

	CA	F/O	F/E
CA		.63*	.59*
F/O			.76***

*p < .01; **p < .001; ***p< .0001

STUDY 3

METHOD

Material and Procedure

This study used the same material and procedure as in study 1. As before, subjects received the 22 scenarios and were asked to sort them into piles of events that involved the same kind of major decision. The critical difference to study 1 concerns the experience level of the participating pilots. The captains in this group of subjects had on average the same years of experience as the first officers in study 1. If this group of captains sorted similarly to the previous, more experienced, sample of captains, then we can conclude that crew function governs feature selection. On the other hand, if years of experience determines what aspects pilots consider for decision making, then this new group of captains should replicate the dimensions observed for the first officers in study 1.

Subjects

Twenty-nine pilots from a different US airline volunteered as subjects. There were 13 captains with an average of 15 years of experience in civil transport. The 16 first officers had on average 10 years of civil aviation experience.

Analyses

Subjects' sortings were analyzed in Multidimensional Scaling and Hierarchical Clustering procedures. Separate analyses were conducted for captains and first officers. The dimensional values obtained for the scenarios with Multidimensional Scaling were used as dependent variables in Multiple Regression analyses. The predictor variables in these analyses were the captains' (or the first officers') ratings of risk, time pressure, problem definition and response determinacy that were collected in study 2.

RESULTS AND DISCUSSION

Only the results of the captains will be included here since the behavior of this group is the relevant contrast to study 1. Multidimensional Scaling and Hierarchical Clustering analyses suggested that a two-dimensional solution accounted best for their sorting judgments. Subsequent Regression analyses yielded dimensional labels similar to the ones observed for the captains in study 1.

The primary dimension for this new group of captains was also “Risk.” The ratings for risk obtained from captains in study 2 were the single best predictor for dimension 1 ($R^2 = .42$; $F(1,20) = 14.65$, $\underline{p} < .01$). The captains' second dimension was best predicted by ratings for response determinacy ($R^2 = .41$; $F(1,20) = 13.63$, $\underline{p} < .01$). Ratings for time pressure picked up an additional 13% of the variance in the scores ($R^2 = .53$; $F(2,19) = 10.7$, $\underline{p} < .001$) indicating that this dimension also entails some evaluation of time pressure.

These findings support the view that pilots focus on situational aspects that are consistent with their crew role. Years of aviation experience per se does not seem to have a significant impact. The captains in study 3 showed preferences similar to their more senior colleagues in study 1 and differed in their judgments from the first officers in study 1 who were their equal in years of aviation experience.

GENERAL CONCLUSIONS

In this series of studies we investigated what situational aspects expert pilots use in making aviation decisions. Our analyses both confirm and expand Orasanu's (1994) model of aviation decision making. They confirm the model with respect to its components. The situational features, “Risk,” “Time Pressure,” and “Response Determinacy” that were discerned in our current work, are also included in the model. However, the model does not sufficiently acknowledge the primacy of risk and time assessment that was evident in our analyses. Our findings suggest that both variables greatly shape what pilots will do. Risk and time pressure may call for an immediate response whether or not the problem was fully understood. Minimal risk levels and time constraints, in contrast, permit additional diagnostic actions or the deliberation of options.

Our analyses also indicate that there is not a single set of features that all pilots exploit. Instead, we found that captains, first officers and flight engineers focused on those aspects during hypothetical decision making that were consistent with the particular responsibilities associated with their crew role. The captains were particularly sensitive to risk and time pressure; i.e., to situational features that are fundamental to aeronautical decision making insofar as they delineate the ground on which decisions can be made. First officers and flight engineers emphasized the response side of the decision making process[2]. They distinguished between situations that provided for only one response and those situations that left options.

These differences between crew positions do not reflect knowledge differences. We found that first officers and flight engineers agreed with their captains in assessments of risk, time pressure, and response determinacy. The differences in feature selection that we observed between pilots, thus reflect role-specific differences in the <u>perception</u> of problem situations.

---

[2] Additional support for this conclusion comes from the sortings of first officers in study 3 whose primary dimension was “Response Determinacy.”

What are the implications of role-specific perceptions for crew training? First we want to emphasize that our research does not warrant any normative training recommendations in the sense of what situational features pilots should or should not use in aviation decision making. However, our studies do suggest that crew training should stress the need for crew members to discuss openly their views on problem situations since diverse perspectives are possible.

A second implication of our studies for crew training concerns pilot upgrading. Given that pilots' perceptions of problem situations depend on their pilot role, then we cannot expect that first officers who upgrade bring a captain's mind-set to decision making. Training accordingly should include instruction in role-consistent approaches to problem situations.

REFERENCES

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology,* 4, 55-81.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 6, 37-75.

Orasanu, J. (1994). Shared problem models and flight crew performance. In N. Johnston, N. McDonald, & R. Fuller (Eds.), *Aviation psychology in practice*. Avebury Technical.