PILOT JUDGMENT: TRAINING AND EVALUATION

DR. RICHARD S. JENSEN

University of Illinois at Urbana-Champaign

JUDGMENT — perhaps the key to longevity in naval aviation. Judgment is obviously critical in the inflight regime where multiple decisions must be made in a timely, correct, and often irreversible fashion. Just as important, and perhaps too often overlooked, is the importance of good judgment on the ground. Knowing when not to fly, planning flights carefully, and realizing personal limitations are examples of good judgment displayed on the ground (Dunn, 1977).

INTRODUCTION

From the beginning of aviation history pilots have been expected to exercise a considerable amount of judgment in the overall task of flying an airplane. However, in recent years, increasing demands in our society for safety, dependability, economy, effectiveness, and reduced energy consumption have increased the complexity of civil and military flying operations magnifying the pressures for good pilot judgment. Furthermore, technological advances that have eased much of the pilot's burden for precise aircraft control have not greatly eased the pilot's decision-making workload. In many cases these advances have only created demands for higher levels of skill, knowledge, and judgment to which few pilots have been trained, and the training costs to prepare them to operate effectively in the changing system are becoming prohibitive.

Flying has developed so rapidly that there has been little time for a serious study of what flying is all about, particularly in terms of how pilots think. Many changes in regulations are expedients designed to solve problems that have already developed. Often solutions to existing problems create new ones, which in turn, are "solved" by new regulations. The problem of training new pilots and retraining current pilots to facilitate the implementation of new procedures and regulations in a mobile but energy limited society is just beginning to be recognized (Roscoe, 1974).

However, if it were merely a matter of teaching flying skills, the training of pilots to operate safely in our complex aviation system would be a much smaller one than it is. Unfortunately, because actual conditions are never quite the same as those used to develop aviation regulations, procedures, and performance limitations, the safety of a given flight also depends upon a significant amount of evaluation and interpretation of existing conditions by the pilot.

For example, the conditions used to develop flight performance values for a particular type of airplane may be ideal including clean airplane surfaces, a new engine, a new propeller, an unrestricted air filter, and a company test pilot. In actual conditions the pilot must compare these book values obtained in ideal conditions with those in which he finds himself. These actual conditions may include a dirty airplane, a slightly used engine, a few marks on the propeller, a slightly dirty air filter, and a less than perfect pilot. He must then evaluate many other conditions such as gross weight, center of gravity, wind, temperature, humidity, altitude, etc. for comparison with those used in the book to determine his expected flight performance. Finally, he must check the present and forecast weather, the terrain, and expected traffic density and compare them with an estimate of his own capability before determining whether or not his planned flight will be safe.

Examples such as these requiring decisions with less than perfect information are available in all areas of flight activity. Furthermore, every decision that the pilot makes is colored by physiological, psychological, and social pressures that are virtually impossible to weigh properly on the spot. For example, just as persons watching a sporting event may "see" an infraction or foul differently depending upon their vantage points and which team they support, a pilot may be influenced to view the weather outlook or his own abilities differently depending on the importance or value he assigns to a given flight. The person's self-image and his need to maintain his external image largely determine how much effect different values or rewards for making a flight will have on his judgment of his ability to make a safe flight (Kogan and Wallach, 1964).

Mental weaknesses in some pilots may cause them to be susceptible to social pressures that result in less than rational pilot judgment. Such irrational pilot judgment is characterized by such unsafe practices as flying under bridges, landing on busy

highways, attempting to land in football stadiums, and flying "formation" on other unsuspecting pilots. Potential sources of social pressure that may lead to these types of activities include peer reactions, fear of failure, censure from superiors or family members, and many others (Janis and Mann, 1977).

Although it may be a difficult task (Fishbein and Ajzen, 1975), pilots must be tested, not only for their knowledge, skill, and rational judgment capabilities, but also for their irrational judgment tendencies as these apply to safe flying. It is apparent from accident statistics that new approaches to pilot training and testing are needed to improve the safety and effectiveness of civil and military pilots.

Training and Testing Effectiveness

An assessment of the effectiveness of current pilot training and testing programs should start with a categorical analysis of training objectives associated with the end product: a pilot licensed to fly under a certain set of regulations. Civilian training objectives may be classified under three sets of behavioral activities as follows:

Procedural Activities

- Communication management
- Navigation management
- Fuel management
- Powerplant management
- Vehicle configuration management
- Display management
- Autopilot management

Perceptual-Motor Activities

- Vehicle control
- Distance, speed, altitude, and clearance judgments
- Hazard detection and avoidance
- Communication
- Geographic orientation

Decisional Activities

- Pilot self-evaluation of skill, knowledge, physical, and psychological condition
- Navigation planning
- Hazard assessment
- Assessment of attention requirements
- Assessment of aircraft and ground system capabilities
- Mission priority adjustment

A useful next step in the examination of the effectiveness of current pilot training and testing programs that may help to identify weaknesses is to analyze general aviation accident data in which pilots were "found to be a contributing cause or factor." Statistics from the National Transportation Safety Board (NTSB) Automated Aircraft Accident and Incident Information System from 1970 through 1974 were used in this analysis (Jensen and Benel, 1977). Pilot cause/factors from the NTSB data were classified into the three behavioral categories given above. Then the total numbers of both fatal and non-fatal accidents during the five-year period were determined for each of these behavioral categories. The results of these analyses are shown in Table 1.

Table 1. Number and percent of the total general aviation accidents in which the pilot is listed as a cause or factor between 1970 and 1974.

	FATAL	NON-FATAL
Procedural	264 (4.6%)	2230 (8.6%)
Perceptual-Motor	2496 (43.8%)	14561 (56.3%)
Decisional	2940 (51.6%)	9087 (35.1%)

Although these satistics of pilot-caused accidents reflect the influence of more factors than just pilot training deficiencies alone, examinations of these data provide valuable indications of possible weaknesses in current programs. For example, a majority of the non-fatal pilot-caused accidents (56.3 percent) were the result of faulty perceptual-motor behavior. The most significant factors here (failure to maintain flying speed and misjudgment of distance, speed, altitude, or clearance) represent one type of pilot judgment. On the other hand, a majority of the fatal pilot-caused accidents (51.6 percent) were the result of faulty decisional behavior, another type of pilot judgment. The most significant factors in this area were the familiar "continued VFR into known adverse weather" and "inadequate preflight planning or preparation."

It is apparent from these accident statistics that both aspects of the deciding function are important to safe flight and possibly suffer from neglect in the present training and testing process. However, because it suffers from greater misunderstanding in aviation circles, pilot judgment as represented by the general decisional activities is the topic of concern in this paper. Although a significant amount of research has been done on this aspect of judgment in recent years (Janis and Mann, 1977) no one has specifically examined this judgment problem faced by the pilot, the flight instructor, and the pilot examiner.

There appear to be three major problems that require solution before major improvements to pilot training and evaluation

can be realized in this area. The first is the establishment of a common definition of judgment as it applies to flying. At present, even though the term is used repeatedly in aviation circles and FAA examiners are required to evaluate candidates on the basis of judgment, no such definition exists.

The second major problem is to determine whether or not pilot judgment can be taught, and if so, how can one best teach it. Because some aspects of pilot judgment are closely akin to personality characteristics, they may be difficult to modify. It may be necessary to use testing and selection procedures to improve aviation safety and effectiveness from these standpoints. Other aspects of pilot judgment are more easily modified through systematic training procedures.

The third major problem is to determine whether or not pilot judgment can be evaluated reliably, meaningfully, and objectively. Because judgment is primarily a mental process, it may be difficult to evaluate in any reliable way. On the other hand, behavioral events frequently have been used to infer mental activity. Although personality tests have proved to be somewhat unreliable, research results using these instruments may be useful in the development of instruments for evaluating and predicting judgmental behavior (Fishbein and Ajzen, 1975).

Judgment Definition

As indicated above, the word judgment has been used to describe two somewhat different mental processes in aviation. Perhaps its most common usage has been to describe the mental activity that takes place at the perceptual-motor level. The second describes the mental activity involved in choosing a course of action from among several alternatives. Obviously, this second usage of the term is similar to the first in that both involve making choices.

However, there is a basic difference. The first refers to highly learned perceptual responses that must be made in a very short time, in some cases continuously. The second, refers to cognitive decisions for which set procedures have not been established or may have been forgotten. Flight instructors have used various terms referring to this type of judgment including "headwork," "thinking ahead," and "staying ahead of the aircraft." Usually, more time is available to evaluate the situation, a larger number of possible courses of action must be considered, and there is a greater degree of uncertainty concerning the existing situation and possible outcomes than is the case in perceptual judgments. For these reasons, cognitive judgments have been the source of greater misunderstanding in pilot training and evaluation.

These two aspects of judgment may be considered as two ends of a continuum based on cognitive complexity and decision time. One such representation is shown in Figure 1. At one end of the continuum, are the common perceptual judgments of distance, altitude, speed, and clearance. These perceptual judgments are less complex in that they involve fewer pieces (frequently one) of fairly accurate information from which responses are determined with highly learned motor behavior. They may require simple responses but frequently call for immediate control movement.

At the other end are what might be called cognitive judgments. As described above, these judgments are very complex in that they usually involve a large number of relevant pieces of highly probabilistic information, they usually require the specification of and choice from among several alternatives, and they are frequently affected by emotions, values, and social pressures. In addition, cognitive judgments usually permit some deliberation before a control response is required. The remainder of this paper is concerned with this aspect of pilot judgment.

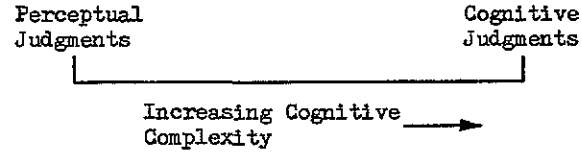

Figure 1. Judgment Continuum Based on Cognitive Complexity and Decision Time

Cognitive Judgment Definition. Considering these factors a candidate definition of cognitive judgment in flying airplanes is:

1. The ability to search for and establish the relevance of all available information regarding a situation, to specify alternative courses of action, and to determine expected outcomes from each alternative.

2. The motivation to choose and authoritatively execute a suitable course of action within the time frame permitted by the situation.

Where:

1. "Suitable" is an alternative consistent with societal norms.

2. "Action" includes no action, some action, or action to seek more information.

The first part of the definition refers to intellectual abilities. It depends upon human capabilities to sense, store, retrieve, and integrate information. This function is what Van Dam (in Jensen and Benel, 1977) calls the "discriminating ability" in professional pilots. In signal detection theory it is called detectability (d'). It is purely rational and could be stated mathematically. If it were possible to separate this part of human judgment from the second part (which it is not), the mind would solve problems in much the same way as a computer. This is not to say that it would be error free. It uses probabilistic information and is dependent upon the amount, type, and accuracy of information stored as well as inherent and learned capabilities to process information.

The second part of the definition refers to motivational tendencies. The emphasis is on the directional aspects of motivation rather than the aspects of motivation dealing with intensity. It says that a part of human judgment is based upon bias factors (costs and payoffs) or tendencies to use less than rational information (defined by society) in choosing courses of action. Society would probably consider the use of any information other than that required to define the safety risk (e.g., monetary gain, gain in self-esteem, adventure seeking, etc.) as less than rational. This part of human judgment is called the response bias (B) in signal detection theory. It is what Van Dam (in Jensen and Benel, 1977) has called the "response pattern" of the professional pilot. If properly developed, this part of human judgment would tend to halt the use of information not directly related to the safety of the flight and to direct the pilot's decision toward the use of rational processes.

JUDGMENT TRAINING

Can Pilot Judgment Be Taught?

The first question to be addressed following the establishment of the definition is whether or not pilot judgment, as defined, can be modified through training. The paucity of judgment training guidelines in pilot training and training research literature leads one to doubt that judgment can be taught. Literature and syllabi commonly used in flight instructor courses contain large sections on how to teach the motor skills of flying but very little on how to teach pilot judgment (see the FAA's Aviation Instructor's Handbook, 1977). The typical private pilot course offers a scattering of judgmental instruction in the areas of weather avoidance and power-plant emergencies but no systematic judgmental training.

However, there is evidence in aviation showing that at least one form of judgmental training, assigning procedures for every conceivable situation that might arise, may be effective. In the military these are referred to as "Boldface" training procedures. Demonstrations by American Airlines (Gibson, 1969) and by Trans World Airlines (Trans World Airlines, 1969) offer convincing support for the conclusion that complex simulators are effective both for the training and testing of pilots using these procedures.

Looking outside the field of aviation one finds other evidence indicating that judgment may be taught. For example, although the theory of signal detection (TSD) was not designed specifically to handle cognitive judgments, many of its methods can be used to explain and perhaps even modify pilot judgment behavior. TSD divides an individual's decision behavior into two components representing his sensitivity (d') and his response criterion or bias (B), roughly corresponding to the two aspects of our judgment definition.

The sensitivity is affected both by the physical value of the stimuli in the situation (signal vs background noise) and the quality of the sensory apparatus of the observer. In cognitive judgment this is the intellective component. On the other hand, the response criterion represents the point in the signal-to-noise distribution at which the observer is willing to say "signal." It is the amount of information, in the presence of noise, needed to tip the decision one way or the other. It is influenced by motivation, knowledge of the signal's probability of occurrence, and the costs and payoffs attendant with a given response. In cognitive judgment the response criterion is the motivative component.

The response criterion can be manipulated through a wide range of values by adjusting probabilities, costs, and payoffs (Birdsall, 1955). We can infer from the vast amount of psychophysical decision data that cognitive judgments can be modified in a similar way. Decision biases, attitudes, risk tendencies, consideration for passenger safety, and pilot motivation can and are being taught by the flight instructor by example, if not by design, at all levels of pilot training. These tendencies are taught, perhaps unconsciously, by the assignment of probabilities, costs, and payoffs to actions of the student by the instructor.

Although TSD says that the sensitivity component is quite stable for a given individual, there is a growing field of research indicating that, if considered as the intellective component of cognitive judgment, sensitivity can be modified as well. For example, attempts have been made to discover the mental processes that are used by expert

judges such as stock brokers, livestock judges, and medical diagnosticians in making their decisions (Shanteau and Phelps, 1977; Slovic, 1969; Anderson, 1969; Hoffman, Slovic, and Rorer, 1968). The hypothesis is that if models of the mental processes used by these experts in decision-making were available, they could be used in training others to use similar processes. In each of the areas studied, judgmental training traditionally occurs over a fairly long apprenticeship program in which the trainee observes the expert make decisions and learns by this observation. However, as in aviation, because of the complexity of the information used to make decisions, observation or even trial and error are inefficient training methods.

The research on the motivative aspect of cognitive judgment also indicates that training can have a beneficial effect. The major research efforts in this area are reported by Janis and Mann (1977). These authors, speaking from a clinical perspective, begin with the assumption that psychological stress is a frequent cause of errors in decision making. They say that stress arises from at least two sources. First, the decision-maker is concerned about the material and social losses he may suffer from whichever course of action he chooses, including the costs of failing to live up to prior commitments. Second, he recognizes that his reputation and self-esteem as a competent decision-maker are at stake. The more severe the anticipated losses, the greater the stress.

Janis and Mann have constructed a "conflict-theory" model of decision-making postulating that the way we resolve a difficult choice is determined by the presence or absence of three conditions: "awareness of risks involved," "hope of finding a better solution," and "time available in which to make the decision." They have developed several clinical procedures to improve decision-making under the titles, "awareness-of-rationalizations," "emotional role playing," "balance sheet," and "outcome psychodrama." They report that these procedures have demonstrated effectiveness in changing decision-making tendencies and in attitude modification.

A Systematic Approach to Training

The need for pilot judgmental training has been established for all levels of flight instruction. Without a systematic judgmental training program, good pilot judgment is acquired by the cautious and the lucky over years of flying experience in many varied situations. Our task as aviation educators using systematic judgmental training techniques should be to compress a lifetime of flying experience into a relatively short training program to instill good pilot judgment into the emerging private or military pilot.

The evidence presented in the preceding section indicates that many aspects of pilot judgment can be taught. The questions that remain are: what approaches should be taken to implement pilot judgmental training and what techniques should be used to evaluate the level of judgment possessed by a pilot or flight student. This section presents a systematic approach to pilot training emphasizing judgmental instruction.

Before proceeding, it is necessary to establish some definitions and constraints involved in this approach. First, training and education, which have been distinguished elsewhere (Glaser, 1962), will be considered equivalent and defined as the "systematic acquisition of skills, rules, concepts, or attitudes that results in improved performance in another environment" (Goldstein, 1974). The "systems approach," a term that has been used and abused in many ways, should, in this training context, emphasize the specification of instructional objectives, precisely controlled learning experiences to achieve these objectives, criteria for performance, feedback within the system, and a recognition of the interaction among system components.

In addition to these, the approach to pilot judgmental training should consider the following constraints: the cost and time required of both student and instructor, the qualifications required of the flight instructor and examiner, and the safety requirements to administer such a program. Finally, although a systems approach is used to develop the training context, the major burden of judgmental training falls directly on the flight instructor. He or she is responsible for the creation and use of innovative situational teaching techniques.

An Instructional Model. A model of an instructional system adapted from one developed by Goldstein (1974) is useful for the establishment of the system context for pilot judgmental training. This model shown in Figure 2, presents five basic interrelated phases in a closed-loop instructional system: assessment, selection, development, training, and evaluation. All five phases are needed to accomplish the goals of a systematic approach to any instructional program. The feedback from the evaluation phase to the assessment phase indicates that an instructional system is never complete. It needs continual adjustment based on the results of the evaluation phase and inputs from the environment.

The assessment phase consists of the establishment of the instructional need and a

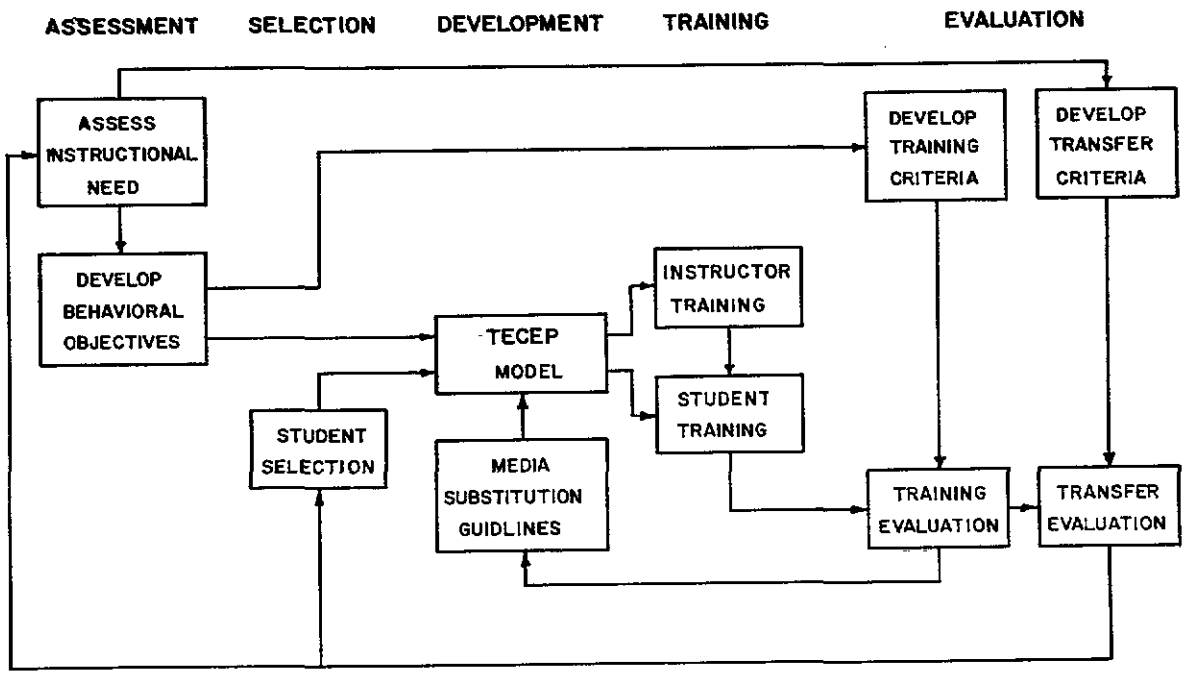

Figure 2. A Model for Pilot Judgmental Training

derivation of behavioral objectives. An assessment of the need for pilot judgmental instruction first requires an anaylsis of the present and future requirements for pilot judgment in the flying environment. Second, this assessment requires an analysis of the judgmental task from the behavioral standpoint. Third, it requires an analysis of human attributes necessary to perform the judgmental tasks. These three analyses provide the basis for the development of judgment behavioral objectives.

Behavioral objectives should specify what the trainee will be able to accomplish when he successfully completes the instructional program. They should also indicate the conditions under which the performance must be maintained and the standards by which the trainee will be evaluated. Thus, they provide direct inputs both into the evaluation phase and the development phase of the instructional model.

The selection phase consists of a program of psychological tests primarily aimed at the identification of persons likely to exhibit irrational judgment behavior during their flying career. For obvious reasons, this phase would be easier to apply in military settings than civilian settings, even though accident statistics have shown that such selection methods might save lives. The necessary instruments for effective application of this phase are yet to be developed.

The development phase consists of the establishment of a training program to achieve the behavioral objectives. The development of this program requires a blend of learning principles and media selection based on the skills, concepts, and attitudes that are to be transferred to the operational flying environment. The learning principles are integrated and matched with appropriate training media in the Training Effectiveness *Cost Effectiveness Prediction* (TECEP) model (Braby, Michelli, Morris, Okraski, 1972).

The training phase consists of both instructor and student training programs. Because judgmental instruction requires the use of situational teaching techniques, instructors will need special training to administer these techniques. The instructor training program will also serve as the source of much of the situational material used in judgmental training.

The evaluation phase consists of the systematic measurement of changes brought about by the training program. Thus, the evaluation phase requires the establishment of measures of success (criteria), based on the behavioral objectives, and the measurement of judgmental behavior both before and after the

training process. Although rarely seen, data from these evaluations are vital to the success of any training program. As Goldstein (p. 23) points out, "...instructional programs are research efforts that must be massaged and treated until the required results are produced."

In pilot judgmental evaluation there also is a requirement for the assessment of judgmental capabilities and tendencies relative to an absolute standard determined by what society expects of pilots. This evaluation would be included as a part of the pilot certification process. Such an evaluation requires a knowledge of societal demands concerning pilot judgmental capabilities, at least qualitative criteria against which to judge the candidate, and an unbiased observation of performance to determine whether or not the candidate meets the criteria.

Some Learning Principles

Because of the common misapplication of some well established learning principles in many training programs, a discussion of these principles as applied to pilot judgmental training is needed. Perhaps the most popular of these is the assumption that the best way to learn an activity is to practice that activity (Gagne, 1962). This assumption is rooted in much of the educational literature and is often identified by the catch-phrase "learning by doing." Gagne points out that it may also be a generalization of the research on the conditioned response in which learning, particularly in animals, appears to have occurred only after a response (practice) has been made.

However, Gagne argues that practice is not an effective training method by itself, even for the acquisition of such motor skills as field gunnery. He says that "instruction about the correct sighting picture for ranging is more effective in bringing about improved performance" than is practice on the task. The point is that training should emphasize the principles and procedures (thought processes) involved, and practice should be directed to take advantage of these principles or take a minor role. If this is the proper emphasis for teaching motor skills, it is even more important in the teaching of judgmental skills which are more highly rooted in thought processes.

A second learning principle that is frequently misapplied in training situations is variously called reinforcement, feedback, or knowledge of results during practice. This principle has been found to be most effective in choice behavior. However, Gagne points out that some manipulations that artificially improve feedback during practice failed to show reliably better transfer to the operational environment and others showed negative transfer. Apparently the form of the feedback is important.

Any beginning flight student will tell you that the usual feedback information such as "you did it right" or "you did it wrong" is almost useless. The time period between trials and feedback may be long, it is often cluttered with interferring information, and the trials themselves are often so complex that the student may learn very little from such a response by his flight instructor. The student really needs to know why he did it right or wrong. He needs to know what rules he should have followed and where he strayed from those rules. Although practice and right/wrong types of feedback may be useful in some training situations, they should be de-emphasized in favor of these "thought" oriented teaching principles in all types of pilot training, but especially in judgmental training.

Judgmental Training Media

Because of the nature of the subject matter to be taught (i.e., attitudes, principles, and motivations), the primary load of pilot judgmental training must be borne by the flight instructor. Practice and conventional self-teaching techniques (e.g., solo flying in a practice area) are highly inefficient methods for imparting these concepts. The following is a discussion of some suggested judgmental training media and techniques that could be applicable to pilot training in civil and military aviation.

Ground School. There are a number of excellent ways that pilot judgment could be taught from the perspective of the conventional ground school. To afford it proper emphasis, it is suggested that judgment should be given a special section of ground school with the same status as meteorology, navigation, and Federal Air Regulations. This section could include lectures and/or discussions of aviation accident senarios in which the pilot was a cause or factor, interactive movies, video tapes, slide presentations requiring student judgmental responses at critical points in flight senarios, and independent study of the principles involved in good pilot judgment.

In addition, this ground school section could include instruction in information integration and subjective probability estimation (Goldberg, 1968). Judgmental behavior in expert judges is characterized by chunking, or the formation of clusters of stimulus attributes and response alternatives, for economy in the thought process. Ground school students could be taught to use these processes in their judgmental activity. The

instructor would show how various types of probabilistic information such as weather forcasts, predicted aircraft system malfunctions, and predicted Air Traffic Control problems should be combined in making flying decisions.

The instructor would teach the student how to "think ahead" or anticipate decisions that might have to be made later resulting from present choices of action. Such anticipation permits the gathering of relevant information under lower levels of stress, when errors are less frequent than later in the flight when time-to-decide may become a error causing factor. This section of ground school could also include decision-making training using procedures suggested by Janis and Mann (1977) such as "balance sheet" and "emotional role playing."

Computer-Assisted Instruction. An instructional technique that holds unusual promise for pilot judgmental training and evaluation is computer-assisted instruction (CAI). The great advantage of these systems is that they can teach principles and then permit the student to participate in decision processes, a highly effective learning technique (Fishbein and Ajzen, 1975). The disadvantage of these systems in the past has been their limited availability and high cost. However, recent advances in technology are making them available at a relatively low cost (Trollip and Ortony, 1977).

Although CAI programs are available in several forms, the dialogue systems that permit student-computer interaction unrestricted by preset response alternatives (Alpert and Bitzer, 1970), show the greatest potential for application to judgmental instruction. These systems depend upon a set of stored algorithms that are used by the computer to construct a great variety of responses to student questions. In addition, student responses are not limited to exact duplicates of prestored expected responses. The program recognizes a variety of student responses and is able to proceed accordingly.

Although practice and feedback are frequently used concepts in CAI programs, these could be augmented by presenting principles and reasons for taking certain courses of action. In judgmental training the student could be presented with a flight situation requiring judgment. He could then be asked to respond by listing all of his alternatives and the factors affecting each. He could even be asked to estimate the probability of success for each alternative.

The computer could then examine the flight experience data on the student (entered previously) and the stored accident statistics from similar circumstances. The computer would then respond with comments on the appropriateness of the student's responses, the alternatives that may have been omitted, and the principles that should have been followed in making the decision. The program could then branch to another problem, the difficulty of which would be based on the level of judgmental capability evidenced by the student's responses to the previous problem.

Complexity, realism, and time constraints could be included in the judgmental task by the addition of a simple flight hand-controller and an airplane symbol with a map on the screen. The purpose of this controller would be to provide indications of progress toward a destination and time available for the decision, not for instruction in flight control.

CAI has many advantages not commonly associated with other instructional systems. The most important of these is individualization of instruction. It can adapt to the specific needs of the individual and interact at his current level of ability (Goldstein, 1974). Second, the unencumbered reinforcement capabilities of CAI are a real benefit to the student. It has no personality or ulterior motives to clash with those of the student. Third, CAI systems do not require the presence of a teacher, although it may be beneficial to have one present for occasional consultation. Fourth, they permit standardization of instruction across a wide area. One central computer could potentially support terminals at every pilot instructional center in the United States at a relatively low cost. No student would be handicapped by a bad instructor who underscores weaknesses in the simulation. Fifth, data gathered from student responses could be stored for as long as necessary for use in updating instructional programs or in evaluating individual pilot judgmental capabilities.

CAI also has a number of limitations that may impact judgmental training. First, very little is known about the effectiveness of the instructional techniques described above. Research is still needed to determine how to program such an instructional system most effectively. Second, large outlays of money would be required for hardware to implement such a program. Third, some users might object to the requirement for communication with the computer via a keyboard, although keyboards are rapidly becoming a part of the pilot's way of life. Fourth, Goldstein (1974) expresses concern for the effects of a machine-oriented learning environment on satisfaction, motivation, and development.

Flight Simulation. Of the various alternatives available for pilot training, it is apparent that, for many operators, flight

simulators may be the most viable for all types of pilot training. The education of pilots for military and airline applications has become increasingly dependent upon ground-based aircraft simulators.

In some ways judgmental training in a simulator environment would be more cumbersome than in ground school or CAI because, at least in current practice, it depends upon the instructor to create the simulated flight situation primarily through verbal communication. Nevertheless, the simulated flight environment provides an additional opportunity to teach judgmental principles, if properly structured, in a somewhat more realistic environment than ground school or CAI can provide.

Probably the best way to begin judgmental training in the simulator is to use the airline approach i.e., teaching procedures that are to be followed in each situation that departs from normal flight. This includes system failure detection as well as establishing courses of action to correct or counter system failures. Principles involved as well as corrective procedures would be taught according to this method, and appropriate judgmental performance measures could be developed.

The simulator instruction could also include the creation, by the flight instructor, of judgment-demanding situations that do not involve the failure of systems. These situations would demand decisions such as whether or not to continue a flight into deteriorating weather, decisions about passenger demands for landing at an unfamiliar alternate airport, decisions about weight and balance considering field conditions, density altitude, etc. In all cases the instructor would ask the student to state several alternatives available to him and also to state which he would choose. These situations could be developed from NTSB accident briefs, and they could be a part of the flight instructor's simulator judgmental instruction package.

Simulator judgmental instruction should be treated as a serious and vital part of the flight student's training. The simulator must be treated as an important training aid just as the airplane and the blackboard are treated. The instructor has the opportunity and responsibility to instill serious, mature flight attitudes in his students by his approach to judgmental training. The simulator provides an outstanding medium for teaching a student good judgment. But the training will only be as valuable as the instructor's approach to simulator instruction is serious.

The Airplane. Of all the media available, the airplane is probably the most difficult to use for direct, systematic judgmental training. The reason is that for the sake of safety, convenience, and cost most judgmental problems must be halted before the student sees the final consequences of his decisions. He frequently must take the instructor at his word that his decision would have resulted in a safe or unsafe situation. However, the airplane offers special opportunities for judgmental instruction because the environment is more realistic, it is more meaningful, and therefore, it is more likely to cause a more permanent behavioral change in the student than other training media.

Everything that has been said about instructor attitudes and approaches to judgmental training applies doubly when actually flying the airplane. Effective judgmental instruction in the airplane requires a consistent, disciplined flight instructor who always follows the rules that the student is expected to follow, or provides a good explanation for why he deviates from them. It also requires that the instructor follow the learning principles stated earlier i.e., that practice and feedback are beneficial only when accompanied by direction and explanations.

Judgmental instructions in the airplane should take the form of simulated situations created by the instructor requiring the use of judgment. Such activities should be interspersed throughout the flight training program. Such instruction is already being done to some extent through training in simulated engine failures, other system failures, and all types of stalls. This training could be expanded to include many of the hypothetical situations discussed above. Portions of such simulated situations could be a part of every instructional flight.

It is the flight instructor's responsibility to teach the student that it is not socially demeaning to refuse to fly or to turn around in the face of deteriorating circumstances. Such situations should be made to occur several times during the student's instruction program in the airplane. Pilots have often said that it is most difficult to turn around the first time. In this regard, it is important to teach the student how to avoid the tremendous social pressure that a group of important passengers can exert. The pilot must be taught to isolate himself from his flight naive passengers in all important decisions.

Finally, often one of the most difficult evaluations a pilot has to make is the self-evaluation of his own skill, knowledge, and judgmental capability relative to a

proposed flight. To assist himself in this regard he should develop a list of personal limitations on flight procedures based on his own capabilities. These limitations must be applicable to all flights regardless of who the passengers are or how much they are willing to pay him to make the flight. They should be invoked during a rational moment, and the pilot's resolution should be strong enough to withstand the enormous social pressure to deviate from them either before or during a given flight.

Situational Emergency Training

The Air Force has begun a research program (Thorpe, Martin, Edwards, and Eddows, 1976) aimed at improving pilot decisional processes during emergency situations. Although the goals of this program are more limited than those of judgmental training described above for civil aviation, the approaches suggested are very similar. The proposed training program being studied, called "Situational Emergency Training" (SET), is designed for the F-15 to replace the traditional "Boldface" procedures of other USAF weapons systems. Although Boldface procedures are effective in many situations where their solutions are applicable, the investigators suggest that there are situations in which these solutions may not apply and such training methods inhibit good judgment in these situations.

SET encourages the development of judgment and centers training around three emergency rules: (a) maintain aircraft control, (b) analyze the situation and take proper action, and (c) land as soon as practical. The underlying concept of SET is situational training. The pilot is taught to discriminate between relevant and irrelevant dimensions of situations which are systematically manipulated in the training program. As pointed out above, this discrimination process is fundamental to good judgment. The authors suggest a scenario development program using instructor training courses as one of the major sources of input.

JUDGMENT EVALUATION

Perhaps the most difficult part of any study of human judgment is the evaluation of performance. The reason is that much of what must be evaluated cannot be observed directly but must be inferred from observation of other related behaviors. From discussions with flight instructors and pilot examining personnel, it is clear that judgment is not being evaluated effectively today (Jensen and Benel, 1977).

Although flight test guides published by the FAA specify that civilian pilots are to be evaluated for their "judgment" capabilities, no definition of judgment is provided. For this evaluation, examiners primarily depend upon the judgment of flight instructors who have the opportunity to examine their student's decision-making capability over a greater variety of circumstances. However, in interviews with flight instructors, only one was found who admitted to having failed a student purely on the basis of poor judgment. Although many said that they could recognize poor judgment, students were failed on the basis of a borderline performance of some other more clearly defined flying maneuver involving skilled performance.

Some ideas for judgment evaluation are offered by Van Dam (in Jensen and Benel, 1977). In his approach, the evaluation begins with psychological and intellegence testing prior to admitting students for flight instruction. Initial impressions from these pretraining examinations are augmented with other subjective indicators of judgment such as "obvious effort and attention to instruction," "relaxation," "division of attention," "response delays," "confidence," "capacity for problen-solving," and "initiative." In later pilot training, evidence of judgment development is seen through an "eagerness to learn or high motivation," "teachability," "adaptability and flexibility," "an intuitive quality in thinking or decision-making," "a pattern of good choices," and "application of margins and allowances."

As indicated in the training system model shown in Figure 2, a vital part of any educational system is an effective evaluation program. Training must be continuously modified in response to the results of these evaluations and progress of the individual students is noted.

The requirements of pilot judgmental evaluation are even broader than these. Society expects pilots to make decisions based on the interests of passengers and property owners. Therefore, judgment must also be evaluated in an absolute sense against this poorly defined scale.

There are three major dimensions along which judgment should be evaluated; each presents a unique problem to the evaluator:

1. The assessment of judgmental capabilities and tendencies prior to flight training.

2. The assessment of the effects of training on pilot judgment.

3. The assessment of the amount of training transferred to the operational flying environment.

Pretraining Evaluation

It is important from the standpoints of both safety and economics to identify persons, prior to flight instruction, who may have difficulty with some aspects of flying judgment. If such individuals could be identified, they could either be discouraged from seeking flight training or their training programs could be modified to offset this deficiency.

Unfortunately, on the basis of psychological testing research to date, the predicted success of such a pretraining evaluation program is not very good. For example, psychologists and others have made many attempts, with little success, to identify a general personality trait known as risk-taking and to link this trait to accident proneness (Shealy, 1974). However, Shealy found that if one were to limit the scope of the test to specific situations, such as down-hill skiing, its predictive validity would be greatly increased. Therefore, it would seem that efforts to develop pretraining pilot judgment prediction tests should not be discouraged by the limited success of the general tests. Instead, efforts should be made to design an aviation specific test with with judgment predictive validity.

Pretraining evaluations of judgment ability in pilot training candidates is a potentially useful adjunct to the entire training and evaluation process. Results from such tests could be used by training management to adapt their programs to emphasize training in areas identified as potentially weak in these tests. Flight instructors could be alerted to possible weaknesses in individual students and adapt their training accordingly.

Tests which could identify risk-taking tendencies (Kogan and Wallach, 1964; Taylor and Dunnette, 1974) and tests which indentify accident proneness (Shaw and Sichel, 1971) are potentially useful in this regard. Situation specific tests, as mentioned above, would be useful in this application and for test development for use in later training as well.

Training Evaluation

The second major dimension along which pilot judgment must be evaluated is an assessment of the amount of change in the pilot's judgment performance that is the result of training. This measure provides an indication of the value of the training program as well as indications of individual student progress.

The development of clearly defined judgmental evaluation criteria presents the greatest challenge to effective evaluation of pilot judgment in all phases of pilot training. To insure that evaluations are made along the same dimensions as the training conducted, the development of these criteria should be based on pre-established behavioral objectives. Judgmental criteria should consist of positive statements of acceptable pilot judgmental behavior for each major area of flight activity. Similar criteria could be developed for every major maneuver taught. These could be graded by the instructor together with evaluations of knowledge and skill each time the maneuvers are attempted.

In pilot training, for each level of pilot experience, certain judgmental hurdles (proficiency levels) could be objectively specified. The instructor, or examiner who evaluates the judgments, would have a range of acceptable performances, also objectively specified. Evaluation of pilot judgment would be a matter of comparing performance against the established criteria in carefully structured situations.

The critical point for judgmental evaluation in a national system is the use of the same criteria by all judges as well as by the pilots themselves. One way to insure standardization of judgmental evaluations is to use a nationwide CAI system to administer tests at specific times during each student's training program. Results of such tests could be used to modify the individual student's training or the training program as a whole.

Transfer Evaluation

The final dimension along which pilot judgment must be evaluated is an assessment of the amount of training that is transferred to the operational flying environment. This means that students who have received special judgmental training are compared with those who have not received such training after both groups have moved into the operational flying environment. The results of this evaluation are used to modify both student selection criteria (or pretraining examinations) and program need assessments.

The criteria for this evaluation are basically the same as those used in training evaluations except that they are more highly influenced by societal demands. Measures of judgmental training transferred can be made in terms of the number of accidents or incidents due to faulty pilot judgment reported within the respective groups.

Operationalizing Judgmental Evaluations

The definition of pilot judgment has two components: discrimination among situational dimensions and response selection. Both components must be evaluated. To operationalize these components for use in any specific

training or testing situation, the evaluator may ask the following questions:

1. For discriminative judgment: Did the student consider all of the alternatives available to him? Did he consider all of the relevant information and assign proper weights to each item? Did he integrate the relevant information efficiently before making his choice?

2. For response selection tendencies: Did the student exhibit any tendency to consider factors other than safety (such as his own self-esteem, adventure, or social pressure) in making his response selection? Did he seem to be highly prone to use semi-relevant factors, such as financial gain or convenience, in situations where safety should have been the primary consideration?

The use of such criteria as these requires more of the evaluator than just an occasional passing glance at the instrument panel. It requires the careful structuring of the situation, perhaps hypothetically, and a careful examination of actions taken by the student. It probably would require a dialogue between the student and the evaluator to establish what the student actually considered in making his choice. Each evaluation must be considered a training device as well, and as such, feedback should be given to the student concerning all aspects of the decision situation known to the evaluator. It is recognized that evaluations of this sort place high demands on the flight instructor. Nevertheless, they seem to be warranted in view of the high number of fatalities caused by faulty judgment, a factor that is hardly being evaluated at all under the present system.

REFERENCES

Alpert, D. and Bitzer, D.L. Advances in computer-based education. Science, 1970, 167, 1582-1590.

Anderson, N.H. Comment on "analysis-of-variance model for assignment of configural cue utilization in clinical judgment." Psychological Bulletin, 1969, 72, 63-65.

Birdsall, T.G. The theory of signal detectability. In H. Quastler (Ed.) Information theory in psychology: Problems and methods. Glencoe, IL: Free Press, 1955, pp 391-402.

Braby, R., Michelli, G.S., Morris, C.L., and Okraski, H.C. Staff study on cost and training effectiveness of proposed training systems. Orlando, FL: Training Analysis and Evaluation Group, NTEC, 1972.

Dunn, R.F. The ultimate penalty for poor judgment. Approach, February 1977.

Federal Aviation Administration. Aviation instructor's handbook. Washington, DC: U.S. Government Printing Office, No. AC 61-14, 1977.

Fishbein, M. and Ajzen, I. Belief, attitude, intention, and behavior: An introduction to theory and research. Reading, MA: Addison-Wesley, 1975.

Gagne, R.M. Military training and principles of learning. American Psychologist, 1962, 17, 83-91.

Gibson, J.H. Optimized flight crew training: A step toward safer operations. Fort Worth, TX: American Airlines, Flight Training Academy, 1969.

Glaser, R. (Ed.) Training research and education. Pitsburg: University of Pittsburg Press, 1962.

Goldberg, L.R. Simple models of simple processes? Some research on clinical judgments. American Psychologist, 1968, 23, 483-496.

Goldstein, I.L. Training: Program development and evaluation. Belmont, CA: Wadaworth, 1974.

Hoffman, P.J., Slovic, P., and Rorer, L.G. An analysis of variance model for the assessment of configural cue utilization in clinical judgment. Psychological Bulletin, 1968, 69, 338-349.

Janis, I. and Mann, L. Decision making: A psychological analysis of conflict, choice, and commitment. New York: Free Press, 1977.

Jensen, R.S. and Benel, R.A. Judgment evaluation and instruction in civil pilot training. Springfield, VA: National Technical Information Service, Final Report FAA-RD-78-24, 1977.

Kogan, N. and Wallach, M. Risk taking: A study in cognition and personality. New York: Holt, Rinehart, and Winston, 1964.

Roscoe, S.N. Man as a precious resource: The enhancement of human effectiveness in flight operations. New York: AIAA Paper No. 74-1296, Nov. 1974.

Shanteau, J. and Phelps, R.H. Judgment and swine: Approaches and issues in applied judgment analysis. In M.F. Kaplan and S. Schwartz (Eds.), Human judgment and decision processes: Applications in problem settings. New York: Academic Press, 1977.

Shaw, L. and Sichel, H.S. Accident proneness: Research in the occurrence, causation, and prevention of road accidents. New York: pergamon,1971.

Shealey, J.E. Risk-taking in skilled task performance. Proceedings of the Eighteenth Annual Meeting of the Human Factors Society. Santa Monica, CA: Human Factors Society, 1974.

Slovic, P. Analyzing the expert judge: A descriptive study of stockbroker's decision processes. Journal of Applied Psychology, 1969, 53, 255-263.

Taylor, R.N. and Dunnette, M.D. Influence of dogmatism, risk-taking propensity, and intelligence on decision-making strategies for a sample of industrial managers. Journal of Applied Psychology, 1074, 59, 420-423.

Thorpe, J.A., Martin, E.L., Edwards, B.J., and Eddowes, E.E. Situational emergency training: F-15 emergency procedures training program. Williams Air Force Base, Arizona, Technical Report No. AFHRL-TR-76-47(1), June 1976.

Trans World Airlines. Flight simulator evaluation. Kansas City, MO: Trans World Airlines, Flight Operations Training Department, 1969.

Trollip, S.R. and Ortony, A. Real-time simulation in computer-assisted instruction. Instructional Science, 1977, 6, 135-149.

ABOUT THE AUTHOR

DR. RICHARD S. JENSEN is on the staff, Department of Psychology, University of Illinois. He serves as principal investigator on a NASA contract which involves development of predictor algorithms and symbology for head-down cockpit displays, is doing a study for the Federal Aviation Administration (FAA) on pilot judgment instruction and evaluation in civil pilot training programs, and acts as a consultant. For seven years, he was a research associate for the Aviation Research Laboratory at University of Illinois. He was program manager on FAA contracts investigating pilot performance on navigation systems. He directed a study of visual landing simulations for pilot training programs and worked on a program sponsored by National Research Development Corporation of the United Kingdom on pilot performance for graphic area navigation. He holds a B.S. degree in physics from University of Wisconsin and the M.S. and Ph.D. in psychology from University of Illinois. He received pilot training at Moody Institute of Aviation and holds an airline transport pilot's certificate with a multiengine rating; a flight instructor's certificate with airplane, instrument, and multiengine ratings; and an aircraft mechanic's certificate with airframe and powerplant ratings. He is a member of the Human Factors Society and Institute of Navigation.