Integrated Framework To Analyze Coordination and Communication Among Aircrew, Air Traffic Control, and Maintenance Personnel

Suzanne M. Dawes

Human error has been cited as a factor in many aviation incidents. Increased automation has not decreased the number of incidents related to human error, but rather has introduced new classes of errors. These errors often result from a lack of coordination and communication among the crew—not only the aircrew but air traffic controllers and ground personnel. It is proposed that an individual’s decision style is one tool that can be used to examine the coordination and communication among these team members. Using an advanced aircraft simulator, test pilots flew a generic flight including takeoff, climb, cruise, approach, and landing under both normal and emergency operations. Decision styles were shown to affect work load ratings, the amount of information used during a segment of flight, and the amount and complexity of written information provided.

Aviation incidents related to human factors are of increasing public concern. Despite improvements in the sophistication and reliability of technology, the percentage of human error–related incidents and accidents has not decreased. As stated in the FAA’s National Plan for Aviation Human Factors, “Human error has been identified as a causal factor in 66% of air carrier accidents, 79% of commuter fatal accidents and 88% of general aviation fatal accidents” (1). Other large-scale accidents—such as Chernobyl, Three Mile Island, Bhopal, Vincennes, Avianca Flight 052 (2), and Dryden—illustrate the “consequences of poor human factors planning in the design and operation of complex systems” (3). Attempts have been made by various agencies such as the FAA, National Aeronautics and Space Administration (NASA), and Department of Defense to address this concern. However, “these efforts have not been organized into an overall plan that addresses the comprehensive nature of human factors issues in the operation and maintenance of all types of aircraft, in air traffic control system operation and maintenance and the interface between the air and the ground” (1).

The objective of this research was to apply an additional technique for the assessment of human performance in aviation. The first step in this process was to develop a conceptual framework of skill-, rule-, and knowledge-based decision making between primary players [aircrew, air traffic controllers (ATCs), and maintenance personnel] (4,5). The second step studied the impact that different decision styles had on the performance of aircrews. Within this framework, the team structure of individuals in the cockpit was examined in detail, as were the effects of different team combinations on performing routine (skill- and rule-based) decision making versus unknown or emergency (knowledge-based) decision making.

WORK LOAD

As the development of advanced aircraft systems continues, it is apparent that these new systems are becoming increasingly complex. The availability of computer-aided imagery and data-processing capabilities has paved the way for the introduction of even more complex and sophisticated hardware. As a result of these technical improvements, the aircrew must process greater amounts of information and make decisions in extremely complex environments. Stresses such as fatigue, cultural variables, cost, schedule constraints, and regulatory constraints combine to produce sustained high work load demand on aircrews.

Problems with mental work load occur routinely in aerospace. Overloading situations are known to have occurred in military combat aircraft, commercial jet aircraft, and air traffic control. Any one of the current measures of work load may not be sufficient to address concerns, such as information management and crew coordination. Automation is one solution that is often pursued. However, automation may not represent an optimum remedy because monitoring is still required, and usually conventional or manual backup must be provided (6). In addition, it has been shown that although automation eliminates certain classes of errors, it can also introduce new classes of errors. These errors are often in the form of not detecting when the human must reenter “the loop,” in some cases resulting in errors whose consequences are more severe than the ones eliminated by the automation (7).

Subjective measures are direct or indirect queries of operators regarding their opinion of work load level involved in a task (8). Reid et al. (9) discuss some of the reasons for using multidimensional subjective measures of work load. Practical reasons for subjective measures as a component of a comprehensive work load test are the relative ease of administration, widespread acceptability, minimal instrumentation, and nonintrusiveness with performance of the primary task.

One highly used method for subjective measurement is the subjective work load assessment technique (SWAT) (10). SWAT was designed specifically to assess human mental work load by asking operators how hard they are working. In order to develop SWAT, researchers defined work load as primarily comprising three dimensions: time load, mental effort load, and psychological stress load. Each of the three dimensions has three levels corresponding roughly to high, medium, and low loading. Each of the three dimensions contributes to work load during performance of a task or a group of tasks. All three factors may or may not be correlated. For example,
one can have many tasks to perform in the time available (high time load) but the tasks may require little concentration (low mental effort). Likewise, one can be anxious and frustrated (high stress) and have plenty of spare time between relatively simple tasks. This rating of workload is based on direct estimate or comparison estimate of the work load experienced at a particular time.

CREW RESOURCE MANAGEMENT

Research on crew coordination and communication, more commonly known as crew resource management (CRM), has indicated that both initial and recurrent training in CRM lead to continuing improvement in crew performance over time (11). Helmreich and Foushee (12) provide the following definition: “CRM includes optimizing not only the person-machine interface and the acquisition of timely, appropriate information, but also interpersonal activities including leadership, effective team formation and maintenance, problem-solving, decision making, and maintaining situation awareness.”

Cockpits have evolved from single-seat aircraft to cockpits with multiple crew members and advanced technology. Additional crew members were initially perceived as backup for the pilot. However, as aircraft design advanced, aircrews were increasingly required to work as a team to maintain effective performance (12). To this end, CRM research has expanded its boundaries beyond the cockpit to include flight attendants, ATCs, maintainers, and ground personnel. As Kanki and Palmer (13) state, “There can be no doubt that operating modern aircraft is a high-stakes profession with lives invested in every flight. It is therefore reasonable to assume that communication plays an important part of this human activity as it does in all others where individuals are trying to accomplish common goals and separate tasks.”

As Helmreich and Foushee (14) state, three categories of variables affect group performance: input, process, and outcomes. Input variables include the individual with the knowledge, skills, and abilities brought to the situation and group variables such as structure, size, cohesiveness, and environmental factors. Process factors reflect the interpersonal and technical coordination found in group interactions. Outcome factors define the dimensions of success or failure of the task conducted. Figure 1 shows the major factors that influence the way groups behave and ultimately the outcome of each flight.

Research on accidents found that crews must cope simultaneously with multiple tasks at the group level. Groups differ in their ability to complete these tasks. Reports from the National Transportation Safety Board implicate crew judgment and decision making in 47 percent of fatal accidents (15, 16).

DEcision STYLES

Cognitive styles are defined as learned thinking habits that act as components of an individual’s personality system. Cognitive style represents an individual’s information-processing model; the way she or he receives, stores, processes, and transmits information (17, 18).

Schroeder et al. developed a human information-processing model that suggests that environmental pressures (or load) systematically affect the complexity of information processing of individuals in an inverted U-shaped function (19, 20). Maximum information use is found under a moderate environmental load, and a decline is seen under overload and underload conditions.

Driver and Streufert (21) and Driver and Mock (22) used this information processing and cognitive style model as the basis for a decision style theory. A decision style is developed along two dimensions: information use and focus. Focus is a continuous dimension ranging from unifocus to multifocus. The unifocus style takes the data provided and applies them to a single solution of a decision alternative, whereas the multifocus style takes the same

FIGURE 1 Effective crew performance (14).
amount of data and integrates them into several outcomes simultaneously. In the model, information use is the amount of information the decision maker seeks in making a decision and is split between two types: satisficers and maximizers. Satisficers use just enough information to generate an answer. Maximizers use as much information as possible to generate a solution. The dimensions of decision style are characterized by five decision styles: decisive, flexible, hierarchic, integrative, and systemic. This relationship is shown in Figure 2.

An individual who uses only enough data to generate a sufficient answer has a decisive decision style. For individuals within this category, once a decision is made, it is final. No attempt is made to reevaluate or review additional data. Individuals using this style are concerned with speed, efficiency, consistency, and achievement of results.

In the flexible decision style, the individual also uses a minimalist approach in seeking just enough data to make a decision. In this case, data are used to generate multiple conclusions that are subject to new data, reevaluation with new data, and the generation of a new solution. Driver and Rowe (17), Driver et al. (23), and Driver (24) state that the flexible style is typically associated with speed, adaptability, and a certain intuitiveness.

The hierarchic style seeks the maximum amount of information available to make a decision. This information is analyzed meticulously or reviewed to create the best solution. Once the solution is obtained, it is implemented with a contingency plan, but is essentially resistant to change. This style is often characterized as rigorous, precise, analytic, and even perfectionist (23).

The fourth style—the integrative decision style—like the hierarchic style, uses a maximum amount of information. However, at the same time, this decision maker generates a number of possible solutions. Synthesis is key to understanding this decision style. The integrative style is viewed highly inventive, emphatic, and cooperative. Individuals who do not use an integrative style often view those who do as indecisive.

A fifth decision style is systemic. This individual appears to embody both a hierarchic and integrative decision style. Initially, this decision maker uses a hierarchic approach exploring all options. However, as additional information is presented, this individual is able to integrate it into earlier information as a hierarchic would. The new information results in a prioritized option list. Systemic individuals appear to be more methodical and careful than those with an integrative style, but more open than those with a hierarchic style.

The use of decision style models has many implications for decision making under real-world, time-pressured operating conditions. In mental task performance, different styles consistently demonstrate distinctly different reactions (e.g., perceived difficulty) to the same task load levels and environmental demands (25–27).

**SKILL-, RULE-, AND KNOWLEDGE-BASED DECISION MAKING**

The lowest level of decision making is skill-based decision making, which tends toward decisions based on learned skills. The next higher level of decision making is rule-based decision making, which is often characterized by a single-response situation based on predefined rules of how the situation should be handled (28–30). Orasanu (15) and Wiener et al. (31) further divide rule-based decision making into two types of decisions. The first category is the go–no go decision. An example would be a rejected takeoff arising from "cargo door lights, runway traffic, compressor stalls and/or overheat lights" (15). The second category is the recognition-primed decision (32). In this situation, the decision maker would first interpret the cues as belonging to a particular event and then select an appropriate response on the basis of experience. A recognition-primed decision would be made in the case of a fuel leak where preestablished parameters must be considered (i.e., how much fuel remains, the rate of fuel loss, and how long the aircraft can continue flying). From this information, the closest appropriate airport must be identified and perhaps an emergency declared (15).

The highest level of decision making is knowledge-based decision making. Knowledge-based behavior occurs when the decision maker finds no preexisting structured procedure for the current incident or when the external circumstances allow deep reasoning about the system configuration and evolving phenomena. The decision maker knows intuitively that it is impossible to have a script for every possible circumstance. A script is defined as the product of individual knowledge, operation experience, operational policies, and applicable procedures. Individuals operating on a knowledge basis typically rely on logical reasoning, intuition, and creativity to generate rules for the existing situation. However, when responding to an unfamiliar emergency, they execute control activities based on logical reasoning.

Historically, when an "event" occurs, pilots analyze the event and use a defined hierarchy or procedure to make a decision. This means

![FIGURE 2 Decision style (23).](image-url)
that both expected (changes in displays due to decreased altitude) and unexpected events (low fuel when the pilot thought the fuel level was higher) undergo the same type of processing. Rasmussen (28) reported that as the degree of familiarity decreases, the need for the crew to work as an integrated team increases. Thus, the pilot must make the transition from skill-based (automatic) and rule-based (procedures) to knowledge-based (unknown or unfamiliar territory) decision making, simultaneously moving from performing individual to team tasks.

Knowledge-based decisions can be made for either well-defined or ill-defined problems (29). Well-defined problems include option selections and scheduling decisions; ill-defined problems include nonprocedural activities and creative problem solving. Crews operating with ill-defined problems often try to “diagnose” what is occurring and may be unable to define the problem exactly. In such cases, there is often no prescribed procedure for identifying or solving the problem nor a script for the crew to follow. Such cases may require external input, for example, from air traffic controllers. Because of the ambiguity of the situation, no one correct or best solution is available to the crew (15).

Attempts have been made to reduce crew decision making by automating systems and establishing standard procedures and checklists that are used to cover anticipated failures or emergencies (7). However, even with such automated systems, increases in decision making often occur as a result of adverse weather conditions, unanticipated events (loss of subsystems), or heavy air traffic. While we can categorize these decisions, in reality, for any given flight situation, crews use a combination of skill-based, rule-based, and knowledge-based decision making (15).

RESULTS

Since the SWAT scale development yielded a Kendall W of 0.7528, a group scale was developed. Crew member decision style is summarized in Table 1.

Ten crews completed the simulator session in which each sortie comprised five segments: takeoff, climb, cruise, approach, and landing. Twenty-five total crews were available during the time allocated for study; however, 15 of the crews completed only some of these five segments of flight because of training needs, subsystem evaluations, and time constraints. Only data for the crews completing all five segments were included in the analysis of results. The average work load values for left and right seaters are given in Figures 3 and 4, respectively. In 100 percent of the cases, the work load increased for knowledge-based scenarios.

The data were analyzed to compare information need as a function of decision style. The number of display changes was used as the measure of information need. This analysis indicated that information maximizers (hierarchic, integrative, or systemic decision style) had a greater number of display segment changes for any segment than did satisficers (decisive and flexible decision style). In all five segments, information maximizer pilots had a greater average number of display changes than did information satisficer pilots. The takeoff and landing segments were both statistically significant. For the right seater, the WSO, information maximizers had a greater number of display changes than WSOs who were information satisficers in three out of the five segments. Multifocus individuals (flexible, integrative, and systemic decision styles) generally had more display changes than unifocus (decisive, hierarchic) individuals. In all segments the pilots who were multifocus had a greater

procedure

Before the beginning of the experiment, all crew members completed a base SWAT sort. In addition, crew members who volunteered completed a decision style questionnaire.

Crew members were randomly assigned to teams. All simulator sessions were observed via the operator communications console located within the same general area as the simulator.

Each crew reported to the flight/mission simulator 1½ hr before the beginning of the sortie. Crew members were met by the evaluation conductor and briefed on the mission and requirements of the evaluation. Upon completion of the sortie, they were asked to participate in a 1-hr debriefing.

Work load ratings were collected throughout the sortie. Delayed SWAT ratings were collected at the debriefing to verify the ratings provided during the flight. In addition, each crew member completed a questionnaire to collect subjective information on the various segments of flight and crew coordination and communication during the sortie. Data were collected on the type and location of display used by each crew member during each flight segment.

Crews were randomly assigned to receive alternative rule-based (R) and knowledge-based (K) segments. Half of the crews were presented R-K-R-K-R segments and half were presented K-R-K-R-K segments.

PROCEDURE

Before the beginning of the experiment, all crew members completed a base SWAT sort. In addition, crew members who volunteered completed a decision style questionnaire.

Crew members were randomly assigned to teams. All simulator sessions were observed via the operator communications console located within the same general area as the simulator.

Each crew reported to the flight/mission simulator 1½ hr before the beginning of the sortie. Crew members were met by the evaluation conductor and briefed on the mission and requirements of the evaluation. Upon completion of the sortie, they were asked to participate in a 1-hr debriefing.

Work load ratings were collected throughout the sortie. Delayed SWAT ratings were collected at the debriefing to verify the ratings provided during the flight. In addition, each crew member completed a questionnaire to collect subjective information on the various segments of flight and crew coordination and communication during the sortie. Data were collected on the type and location of display used by each crew member during each flight segment.

Crews were randomly assigned to receive alternative rule-based (R) and knowledge-based (K) segments. Half of the crews were presented R-K-R-K-R segments and half were presented K-R-K-R-K segments.

RESULTS

Since the SWAT scale development yielded a Kendall W of 0.7528, a group scale was developed. Crew member decision style is summarized in Table 1.

Ten crews completed the simulator session in which each sortie comprised five segments: takeoff, climb, cruise, approach, and landing. Twenty-five total crews were available during the time allocated for study; however, 15 of the crews completed only some of these five segments of flight because of training needs, subsystem evaluations, and time constraints. Only data for the crews completing all five segments were included in the analysis of results. The average work load values for left and right seaters are given in Figures 3 and 4, respectively. In 100 percent of the cases, the work load increased for knowledge-based scenarios.

The data were analyzed to compare information need as a function of decision style. The number of display changes was used as the measure of information need. This analysis indicated that information maximizers (hierarchic, integrative, or systemic decision style) had a greater number of display segment changes for any segment than did satisficers (decisive and flexible decision style). In all five segments, information maximizer pilots had a greater average number of display changes than did information satisficer pilots. The takeoff and landing segments were both statistically significant. For the right seater, the WSO, information maximizers had a greater number of display changes than WSOs who were information satisficers in three out of the five segments. Multifocus individuals (flexible, integrative, and systemic decision styles) generally had more display changes than unifocus (decisive, hierarchic) individuals. In all segments the pilots who were multifocus had a greater

subject

Crew members consisted of qualified flight crew personnel who had completed basic aircraft system training. Each crew member was properly attired with standard-issue flight equipment.

apparatus

A high-fidelity six-degree-of-freedom motion base flight mission simulator with a CT-5A visual system for simulation of advanced aircraft was used during this test. The simulator cockpit had a two-person configuration. The left-seat person was primarily responsible for flying the aircraft, whereas the right-seat person was responsible for navigation, threats, and weapons delivery.
number of display changes than those who were unifocus. The differences were statistically significant for all segments with the exception of landing. In three out of the five segments, multifocus WSOs had a greater number of display changes than their unifocus counterparts. Table 2 summarizes the average number of display changes.

No significant differences were found between the number of comments or number of display changes by segment type (rule or knowledge based). In 80% of the segments for pilots and 60 percent of the segments for WSOs, the knowledge-based scenarios had a higher number of comments and display changes than did the rule based. The only segment in which there was a statistically significant difference was the case of the cruise knowledge-based segment for the WSOs. This was due in part to the selection of the knowledge-based scenario. That is, the task selected at the time of the study was not the final procedure for completing the task, thus the increased number of comments.

The number of comments made on the postflight questionnaire was also different as a function of individual decision style. In 80 percent of the segments for pilots and 60 percent of the segments for WSOs, multifocus individuals made a greater number of comments than their unifocus counterparts. Information maximizers also provided more written comments than did information satisfiers. The comparison of written comments is provided in Table 3.

The types of comments made by each group are of particular interest. In general, the comments of those who were unifocus were related directly to the tasks completed during the segment and, in many cases, attempted to explain that the work load was probably lower than they reported. In contrast, those with a multifocus orientation discussed and provided a much more exacting and lengthy discussion. A sampling of comments is presented in Figure 5.

Another difference was that those with a unifocus orientation never discussed segments for which they had provided a low work load rating. However, those who were multifocused provided comments even when the work load was rated satisfactory. A sampling of these comments is provided in Figure 6.

Differences in work load ratings were found between individuals who were unifocused and multifocused. In 80 percent of the segments, multifocus individuals rated work load higher than did unifocus crew members for both rule- and knowledge-based tasks. No significant differences were found between the two groups for rule-based tasks; however, significant differences were found for takeoff, climb, and cruise for the WSOs. A comparison of the work-load ratings of the two groups is given in Figure 7.

CONCLUSION AND FUTURE RESEARCH

The results of this study indicate that the decision style of an individual plays a role in experiencing and thus rating the mental work...
load associated with various tasks. The differences in the subjective rating of unifocus versus multifocus styles arose from the crew members’ perceived difficulty with the mental tasks. Perceived difficulty of a strictly cognitive task is a partial function of the “subjective complexity” of the subject and is also influenced by the components of the environmental load imposed on the crew member. The crew members—depending on their decision styles—had a different perception of the environmental load, then the different style-dependent subjective ratings were predictable. For instance, the unifocus crew members (decisive and hierarchic) had a consistent increase from rule- to knowledge-based tasks, whereas multifocus crew members (flexible, integrative, and systemic) had a larger increase from rule to knowledge based. In both the rule- and knowledge-based segments, multifocus crew members reported higher work load ratings than did unifocus crew members. These findings were supported by an increase in the number of display changes and number and complexity of written comments for multifocus crew members.

These findings are important for those involved in the design and operation of advanced aircraft. Using unifocus individuals to evaluate the initial design of these systems may result in higher work load ratings than would be obtained by examining the subject pool as a group. Multifocus individuals also provide a level of detail in terms of written comments that is far more useful than their unifocus counterparts, whose comments focus solely on the task as currently designed. The use of the multifocus crew members’ “what if”

<table>
<thead>
<tr>
<th>TABLE 2  Average Number of Display Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Segment</strong></td>
</tr>
<tr>
<td>Approach</td>
</tr>
<tr>
<td>Climb</td>
</tr>
<tr>
<td>Cruise</td>
</tr>
<tr>
<td>Approach</td>
</tr>
<tr>
<td>Landing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 3  Average Number of Written Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Segment</strong></td>
</tr>
<tr>
<td>Approach</td>
</tr>
<tr>
<td>Climb</td>
</tr>
<tr>
<td>Cruise</td>
</tr>
<tr>
<td>Approach</td>
</tr>
<tr>
<td>Landing</td>
</tr>
</tbody>
</table>
Unifocus Comments

With the a/r only 30 minutes from takeoff the workload here is probably artificially high.

Simulator lighting and visual presentation increased workload over normal daytime levels.

Multifocus Comments

Time control was somewhat confusing due to the implementation of xxx but the early late indication on display is helpful.

The display is helpful but several button pushes are necessary to keep up to speed on the threat in relation to your flight plan. Changing ranges, declutter schemes and updating display is time consuming while navigating through the threat.

Attention to controls and displays increases due to the transition. All displays have lots of information to cross check.

FIGURE 5 Unifocus versus multifocus comments.

or "what else" questions provides valuable information to designers of these systems before their operational use. Whereas the workload ratings are higher for multifocus crew members, their increased use of the information available (especially for integrative and systemic) appears to provide the crew member with more of the information that oftentimes is needed to diagnose the situation. Whereas there were no deviations from established procedures in this study, it is hypothesized that in future studies with knowledge-based scenarios that evolve over the duration of an entire flight compared with one segment of flight, we will see differences in performance from the different decision styles. Another area that needs further investigation is the specific verbal communication patterns between the different combinations of crews. Observation indicates that the more multifocus and information maximizing a crew member, the more likely he or she will be to both seek and share information. The ability of a crew member to seek necessary and sufficient information, both written and verbal, is critical to maintaining safety of flight.

ACKNOWLEDGMENTS

This research was financially supported by the Graduate Research Award Program, sponsored by the FAA and administered by TRB. The author would like to thank Larry L. Jenney of TRB, Joseph A. Breen of TRB, Robert L. Helmreich of the NASA/University of Texas/FAA Aerospace Crew Research Project, and Gerald S. McDougall of Southeast Missouri State University for their com-

Maintaining envelope is intensive and normal for this phase of flight.

Climb with autopilot is easy to maintain.

The early/late indication on the display is helpful.

Information is good and available during this phase of flight.

FIGURE 6 Multifocus comments with low workload ratings.

FIGURE 7 WSO—unifocus versus multifocus workload ratings.
ments. The support of the Institute of Safety and Systems Management, University of Southern California, has been invaluable, especially the support and encouragement of William Petak and Najmedim Meshkati. The author is also grateful to the Northrop Grumman Corporation, specifically the managers of the System Safety and Human Factors Department, Gerald R. Doria and James R. Francis, for their assistance in completing this research.

REFERENCES