NASA Technical Paper 3528

Cognitive Models of Pilot Categorization and Prioritization of Flight-Deck Information

Jon E. Jonsson
McDonnell Douglas Aerospace-West • Long Beach, California

Wendell R. Ricks
Langley Research Center • Hampton, Virginia

August 1995

Contents

Summary

Automated systems on commercial aircraft flight decks have dramatically increased in number over the past decade and pilots now regularly interact and share tasks with these systems. Consequently, this increased interaction has led human factors researchers to focus more attention on the pilot's cognitive processing of flight-deck information and the pilot's overall mental model of the information flow that occurs on the flight deck. The cognitive activities of categorization and prioritization particularly interest researchers because these activities are important and pervasive in managing and processing flight-deck information.

The experiment reported herein investigated how pilots categorize and prioritize information typically available during flight. Fifty-two commercial airline pilots participated in tasks that required them to provide similarity ratings for pairs of flight-deck information and then prioritize this information under two separate contexts. Such results can be expressed with either spatial or clustering representations. For spatial representations, the dimensions may be considered as representing cognitive factors that individuals use to process the information. For clustering representations, the clusters into which information falls represent categories pilots have for the data.

The results suggested three cognitive dimensions that pilots use in categorizing flight-deck information. These dimensions included

1. The flight function that the information is designed to support

2. The strategic or tactical nature of the information

3. The frequency of information referral

The results also suggested four specific high-level categories that pilots use in categorizing flight-deck information. These information categories included

1. Aviation

2. Navigation

3. Communication

4. Systems administration

Although these high-level categories are sufficient for classifying the majority of flight-deck information, it was found that additional differentiation can be made in the aviation and the navigation categories.

The four information categories were also prominent when analyzing how the pilots prioritized flight-deck information. Along the primary dimensions in which pilots prioritized information, elements seemed to cluster according to the four high-level functions they support. Finally, when asked to prioritize this flight-deck information, it was discovered that there was both a high degree of consensus for the importance of the same information across different flight situations and few individual differences among pilots in their prioritizations.

Introduction

The increase of automation on modern commercial aircraft has made it more challenging for flight-deck designers to support the flight crew for two principal reasons.

First, automation has shifted the emphasis on flight crew tasks from physical actions to the cognitive processes required to accomplish the tasks (Dornheim 1992; Ricks, Jonsson, and Rogers 1994; Ricks et al. 1991). Two examples are representative of this shift. A transcontinental flight can be completely programmed at the departure gate with the flight management computer rather than having the crew make sequential changes with the mode control panel as the flight progresses, or now almost never done, handflying the aircraft for the entire flight. On the McDonnell Douglas MD-11, reconfiguration occurs for some failure states without any pilot intervention.

Pilot performance in both of these examples involves monitoring the systems and determining whether they are functioning correctly. Because of these technological advances, designers need to consider issues such as how the pilot will monitor the flight progress, the crew's understanding of the flight management system, the information required to support the pilots, and the cognitive processing and mental models of the information flow of the pilots. The consideration of pilot cognition has thus taken on an added degree of importance in the design of modern flight decks. The development of valid and reliable methods for addressing these issues will therefore become an important component from the designer's perspective.

Second, designers are challenged because of the dramatic increase in the information that is or will be presented to pilots (Braune, Hofer, and Dresel 1991; Ricks, Jonsson, and Rogers 1994; Ricks et al. 1991). Part of this challenge will be to include new sources of information that are designed to safely increase traffic flow (such as data links and global positioning systems), while another part will be to include technological advances in the way information can be stored and displayed to the crew.

The glass cockpit has revolutionized display technology by allowing the presentation of multiple types of data on the same display. This presentation contrasts sharply with earlier flight decks where electromechanical

displays were dedicated to the presentation of one type of aircraft information. Proposed information sources such as an electronic library and an onboard maintenance systems may provide the crew with yet more data. Determining what information to provide and when to provide it for a variety of tasks and functions will become increasingly important.

In addition to the trends previously mentioned, the very fact that pilots share flight-deck functions with automation adds to the complexity of analyzing crew tasks. While automation frees the crew from many routine tasks, it implicitly adds another requirement because shared functions require the crew to have some understanding of how the automation works, what it is doing, and when it is doing it. Additionally, the crew must be able to ascertain when the automation may be incorrectly performing a given task and must be prepared to take over should the automation fail. In fact, recent research in this area suggests that, when evaluating performance effects of automation on the human operator, issues such as situation awareness and the pilot's mental model of the automation need to be considered (Sarter and Woods 1994, Wiener 1989). Thus, the pilot's interaction with flight-deck automation implicitly adds another level of information by requiring the pilot to have both an awareness and an understanding of how the automation operates.

These concerns, frequently referred to as flight-deck information management issues, suggest greater attention will need to be given to the pilot's processing and use of information. In the development of earlier flight decks, human factors engineers principally concerned themselves with form and fit issues (such as whether a control was reachable, a display was readable, or the amount of strength required to actuate a control). These developments suggest that more attention will need to be given to the information flow between the pilot and the flight-deck systems in conjunction with the cognitive processes pilots use to process flight-deck information.

While most of the previous human factors research concerned features that were directly observable and amenable to traditional measurement techniques, cognitive factors lack many of these features. For this reason, rigorous, empirical methods for determining the cognitive processes of the flight crew are required. The research program presented herein represents a first step toward the development of measurement techniques that could be used as part of an early flight-deck design process for assessing the cognitive task loads that are associated with information processing. The availability of such measurement techniques should allow for flight-deck designs that increase the pilot's situational aware-

ness and lead to an appropriate cognitive workload, thereby resulting in a more effective flight deck.

The project reported herein addresses four principal research issues. First is an empirical investigation of how pilots cognitively process flight-deck information. Of particular interest is the empirical identification of both conscious and unconscious cognitive processes. These results will help determine the mental representation pilots have for flight-deck information, which is the initial step for developing cognitive measures of crew performance. Such measures would allow researchers to examine how pilots process the information they use to perform tasks and, if valid measurement techniques can be developed, to determine whether the crew is subject to over or underloading with respect to their cognitive processing abilities.

The second research issue concerns how pilots prioritize flight-deck information. The objectives of this issue are to determine whether a common underlying strategy exists that pilots use to prioritize flight-deck information and whether a discernible classification scheme exists, that will allow the development of a prioritization system by category.

A third research issue is the consistency of cognitive processing among pilots. That is, are pilots using similar cognitive processes for managing flight-deck information. This issue is critical. If significant differences in the way pilots process and categorize information exist, it would be problematic to extend these findings to other areas.

The fourth and final research issue examines whether the cognitive processes used for flight-deck information are invariant across different contexts. This question is concerned with whether the cognitive processes pilots use change as a function of flight context.

Abbreviations

ANOVA	analysis of variance
ATC	air traffic control
ATIS	Automatic Terminal Information Service
CDU	control and display unit
CRT	cathode ray tube
DME	distance measuring equipment
EPR	engine pressure ratio
FMS	flight management system
F.O.	first officer
ILS	instrument landing system
INDSCAL	individual differences scaling

IRS	inertial reference system
L/R	left/right
MAU	multiattribute utility
MCDU	multifunction control and display unit
MDPREF	multidimensional preference scaling
MDS	multidimensional scaling
ND	navigational display
PCPREF	personal computer program for MDPREF
PFD	primary flight display
SAS	statistical analysis system
TCAS	traffic alert and collision avoidance system
VHF	very high frequency
VOR	VHF omnidirectional range

Background

The concerns presented in the Introduction, frequently referred to as flight-deck information management issues, suggest that greater attention will need to be given to how pilots use and process information. In so doing, researchers should study cognitive processes routinely engaged in by flight crews. The categorization and prioritization of flight-deck information represent two such processes. Because of their ubiquitous nature, it seems appropriate to establish, in an empirical fashion, how pilots categorize flight-deck information and how they judge the relative importance of that information. It would be useful to review briefly both areas from theoretical and applied perspectives.

Categorization of Information

Humans are able to process and use large amounts of information that are presented in everyday life because they can organize the information by relating it to prior experiences (thereby making it familiar). This organizational process draws heavily on an individual's long-term memory and is usually referred to as *categorization*. Such grouping considerably simplifies the individual's processing of new information, not only by providing it with a slot in which to reside, but also by providing some general characteristics of the data once it has been placed. This process, along with its benefits, has been nicely summarized by Glass and Holyoak (1986):

> Categorization is a fundamental cognitive process because every experience is in some sense unique. For example, no two apples are entirely alike. However, if each experience were given a unique mental representation, we would be quickly overwhelmed by the sheer complexity, and we could not apply what we had already

learned to deal with new situations. . . . A category system allows us to derive further information about an object that has been assigned to a category. For example, if you have categorized some object as an apple, you can infer how it is expected to taste, that it has a core, and that it can be used to fill a pie (p. 149).

Thus, categorization allows individuals to infer a great deal about a new source of information, because it succinctly and implicitly describes characteristics of the information. Theories of how individuals categorize information are matters of great debate in the cognitive science community. Currently, there are two principal theories of categorization, one employing prototypes, while the other is based on the concept of exemplars. While it is not the intent of this research effort to experimentally differentiate between the two theories, it is important to briefly review each theory to gain some understanding of categorization models of cognition prior to discussing potential applications for evaluating pilot cognition. For a detailed analysis of categorization models, see Ashby and Maddox (1993).

Prototype explanations are predicated on the notion that for any category, individuals develop a mental instance, which is most representative of the category. This representation is referred to as the prototype. New objects are compared against the prototype of each category. If a new item is sufficiently similar to the prototype, the item is assigned to that category. For example, when a person thinks of an aircraft, a prototype theorist would argue that the individual has a mental representation of what most aircraft look like. Mathematically, this typicality can be expressed as an averaging process of all features that compose an aircraft (such as wings, tail, and flight deck). Unusual aircraft, such as the Northrop-Grumman B-2 and the Beechcraft Starship, are viewed as such *precisely because* the prototype is violated.

Exemplar models of categorization differ from prototype explanations, in that the new object is compared with those individual items that compose a category. The membership of an object in a category is defined by selecting that category whose instances differ the least from the new item. This process essentially takes the form of satisfying a minimization function, whereby an object to be categorized is compared on all relevant features to each object in a category. The category selection rule would be to select the category whose members differ the *least* from the object to be classified. With this theory, features of an unusual aircraft, such as the B-2, would be compared to various categories of things which fly (such as airplanes, helicopters, airships, and rockets). The category selection would be based on an analysis of the features of the B-2 and other members of the

category. The B-2 would be assigned to the category whose individual members had features that were the least different from it.

Note that the principal difference between the two theories is that exemplar models compare the item with individual objects within the category, whereas prototype models compare the item with the prototype of the category.

Both models have yielded results consistent with the notion that the presentation of information congruent with an individual's mental categorization scheme leads to superior performance as measured by accuracy and response time. For example, Posner and Keele (1970) studied classification performance for visual random dot patterns that had been memorized by subjects. The memorized patterns were variants of prototype patterns that were constructed by distorting the prototype by randomly moving the dots. Subjects were then tested by using the memorized patterns, the prototypes from which the patterns were constructed, and two variants of the prototypes that the subjects had not seen. As expected, when tested on classification performance, individuals gave the fastest and most accurate responses to the stimuli they had previously memorized. Of particular interest was that as distortion from the original prototype increased, more errors and longer response times were observed.

Similar results have been found for meaningful stimuli such as words and figures (Reed 1972; Rips, Shoben, and Smith 1973). For studies involving semantic stimuli, subjects will respond quicker and/or more accurately to instances that are closest to the category. For a concrete, everyday example, individuals will take longer to correctly judge that a chicken is a bird, than a robin is a bird. This longer judgement time is due to the fact that for the majority of individuals, "robin" represents a more typical instance of a bird than does "chicken."

The results of these experiments have demonstrated three major conclusions. First, inclusion of superordinate category terms allows researchers to see which stimuli most closely cluster around them. In the above example, "robin" and "bluejay" were much closer to the superordinate term "bird" than were "goose" or "duck." Second, the time to make category judgments (e.g., is this instance a member of category x or y?) is directly related to the psychological distance subjects have for the given set of stimuli. Psychological distance can be measured quantitatively by using psychometric scaling techniques. These techniques will be discussed subsequently. The farther apart a pair of items was in the psychological space, the longer it took subjects to make category judgments. In the above example with birds, Rips, Shoben, and Smith (1973) discovered that, when analyzing these data with techniques that allowed investigators to exam-ine their spatial positioning (thus indicating their psychological relatedness), "robin" was indeed closer than "chicken" to the category term "bird." Third, the psychometric techniques allow researchers to identify the cognitive processes individuals use in categorizing and making judgments.

Just as categorization is a pervasive function in everyday life, it frequently occurs on the flight deck. It is possible that pilots principally use two characteristics to categorize and manage flight-deck information. First, both the source (where the information is coming from) and the destination (where the information is displayed or used on the flight deck) provide the pilot with initial cues about content of the information. For example, the flight crew can initially make some inferences about a message from its source (e.g., ATC, dispatch, or ATIS) prior to interpretation. From the pilot's perspective, destination or location is probably most critical because where information is displayed usually indicates the type of information. For example, in a glass cockpit, assuming all CRT's are operative, pilots know what information they can expect to see when they view any of the displays. However, because display technology allows for the presentation of multiple types and layers of information, location may no longer be as predictive as it has been historically.

Second, the functions that the information support will provide an introduction to how the information will be categorized. Traditionally, pilots are trained to aviate (control of the aircraft), navigate (location and destination), and communicate (communicate intentions to ATC). This training provides one high-level framework for categorizing, although further subdivisions may be possible and even desirable. This issue is a major concern of the research presented in this paper.

As automated information management aids are introduced on the flight deck, the information discussed previously suggests that the processing and the categorization of flight-deck information should be consistent with the pilot's cognitive model of flight-deck information categories.

Prioritization of Information

Prioritization refers to the process by which information or actions are ordered along some dimension or by a given attribute. Single or multiple dimensions and/or attributes may exist, depending on the characteristics of the information and the actions. For example, logical dependence, wherein one source of information is required before the next piece of information can effectively be used, could be considered an attribute. In this case, the first source of information acquired would initially be said to have a higher priority than the

subsequent information. While it is important to recognize in any prioritization scheme, logical dependencies are rather trivial. Of considerably more interest are prioritization dimensions that involve relative attributes or dimensions that change with situations and differ among people. This multidimensional and dynamic nature of information is of primary interest for the current research plan. As with categorization, such dimensions and attributes can be conceptualized as cognitive variables that people use to prioritize the information they use to perform some function.

Prioritization is usually considered as an issue within the context of decision theory. With this perspective, new information arrives and the user of this information decides how to prioritize it in the context of the current state. From a decision theory context, this information can be analyzed as a specialized, expected utility problem that is expressed through multiattribute utility (MAU) models (Edwards 1987).

In the MAU framework, an attempt is made to determine what variables the decision maker considers, the relative weight given to these variables, and the utilities for the expected outcomes. As Edwards (1987) states

> Formal decision theory assumes that the (individual) decision maker has a set of *values*, and chooses acts that, as he or she sees it, will best serve them. . . . The consequences of each act, or of each act-event-act-event . . . sequence, are called *outcomes*. Each outcome is conceived of as having a subjective value or *utility*. (pp. 1063–1064)

Once the attributes have been identified, a determination is made of how much weight the decision maker wishes to give each of them. For example, in purchasing a car, a new car buyer may consider reliability, safety, and cost as highly relevant attributes, while color, style, and gas mileage may be considered less relevant. For an individual making a purchase, each of these variables can be ordered in terms of its salience for making a purchasing decision.

Analogously, pilots must prioritize the attributes of new, incoming information relative to that of information currently available and the state of the aircraft. Identification of how pilots normally process data, with emphasis on what attributes are considered particularly salient, and the categories into which different data fall, is a prerequisite for understanding how pilots prioritize information.

Currently, pilots are responsible for prioritizing flight-deck information. In some cases (principally emergencies) task and information prioritization is done for the crew by aircraft system controllers or emergency

checklists. For nominal information, however, pilots have aids to help them determine what information is required and when it is needed. Such aids are represented in the form of

1. Flight operations manuals

2. Procedural checklists

3. Training procedures and company policy

4. Operational demands

In general, high-level principles govern prioritization. Information required to maintain a safe flight will receive a high priority value. Likewise, because some activities must be performed in limited amounts of time, information required to support a time critical activity will be obtained prior to information necessary for less time critical activities. Finally, information can be prioritized on its logical dependency—one piece of information is necessary before another piece of information can be used (e.g., one needs to know the ATIS frequency before gaining access to that information).

Availability of data must also be considered when discussing prioritization. Flight-deck information may be continuously available (a dedicated display) or the pilot may have to engage in a series of actions to obtain it (a number of button presses on the FMS CDU). Alternatively, information may be provided to the crew without a specific request having been made. Primary examples of this include

1. ATC communications

2. Warnings, alerts, and cautions

3. Sources such as data link and intelligent flight-deck aids

Information competing for available cognitive processing resources underscores the importance of understanding pilot expectancy and mental models of information flow on the flight deck.

Psychometric Tasks and Analyses

Recent developments in cognitive research have demonstrated the usefulness of psychometric techniques in representing human knowledge and information processing (Ashby 1992; Nosofsky 1984, 1986, and 1992; Shepard 1987). One method, MDS, calculates a spatial representation among stimuli by using some measure of how the stimuli are related to one another. This spatial representation presents the objects in an *n*-dimensional space, with items similar to one another lying close together, while dissimilar items lie farther apart in the space. Another method, cluster analysis also uses a measure of stimulus similarity that identifies items closely

associated with one another, groups them, and provides a hierarchical representation of the stimuli, thereby allowing the investigator to examine the representation for obvious or intuitive categories. Objects placed in the same cluster are more similar to one another than objects placed in different clusters.

For both methods, when the measures for stimuli relatedness are elicited from human observers, the representation is said to be cognitive or perceptual. Of particular interest to researchers is the potential for cognitive interpretations of the dimensions or clusters emerging from such analyses. Although some of the most sophisticated work in this area has been done in the last 10 years, MDS had been used as early as 1973 by Rips, Shoben, and Smith.

The assumptions underlying the use of scaling and clustering in cognitive experiments are, in general, quite similar. First, the scaling solution for a given set of stimuli is said to be a cognitive representation for how individuals view the relationship among those items. For flight-deck information, this would be a representation of how related or similar the flight-deck information is perceived as being. Second, dimensions extracted from the scaling analyses can be considered as representing the salient cognitive dimensions along which individuals process information. In the present case, such dimensions would reflect those variables pilots use to process flight-deck information.

Categories or clusters of information tell researchers how pilots define category membership. In the aviation domain, three of the most frequently mentioned high-level categories are aviate, navigate, and communicate. Of interest here is whether these categories are sufficient to describe all the information pilots use, or whether a finer distinction may be more appropriate.

The potential for discovering the way in which pilots process flight-deck information should not be considered merely descriptive. Rather, dimensions found through scaling analyses and categories that emerge from clustering could be used in a predictive fashion for evaluating a pilot's cognitive processing under different flight scenarios and new flight-deck designs. Because the psychometric techniques used herein may prove useful in representing the pilot's cognitive processing of flight-deck information, it might be possible to use this data in constructing models for evaluating the cognitive demands placed on pilots as a result of the tasks they are performing. High cognitive task loads could indicate a situation where the crew may be unable to adequately perform a task; therefore, consideration should be given to modifying the design and/or the operational procedure. This consideration could prove especially valuable for evaluating flight decks early in the design process. If

such valid measurement scales can be developed based on the results of the study reported herein, this information could be used early in the flight-deck design process to determine both the potential for information overload in certain flight phases or the potential for tasks to compete for the same set of cognitive resources. Such applications are discussed in appendix A.

Method

The experiment described in this paper was conducted at Langley Research Center. Details concerning the subjects and the procedures are given in the following sections.

Subjects

Fifty-eight pilots participated in this experiment. Of these, six subjects were eliminated from the subsequent data analysis. Three of these six were removed because all commercial airline time was spent as a flight engineer, two subjects had no commercial airline experience, and one pilot indicated he had given erroneous responses because of a misinterpretation of one of the stimuli. The remaining 52 subjects were all commercial airline pilots who were currently flying or recently retired. The mean total flying time for these subjects was 11435 hr with a range of 3000 to 27000 hr. Average commercial experience was 15 years with a range of 1.5 to 33 years. Of the 52 subjects, all but 1 were male. Eight airlines were represented.

Stimuli

Experimental stimuli were obtained from an unpublished information analysis performed on the Boeing 747-400 by John Groce of Boeing Commercial Aircraft Co.

This analysis consisted of determining all information elements on the 747-400, each with its associated source, destination, display modality, message management schemes, and control modality. (An information element represents a discrete unit of data available to the pilot, usually associated with a specific display or interface.) "True Heading," for example, is an information element associated with the navigation display.

This entire list consisted of 396 information elements. To examine information elements applicable to all modern commercial aircraft, terms specific to the 747-400 were eliminated by having three commercial pilots (none of whom flew the 747) go through the list of information elements and identify terms unfamiliar to them. These items were subsequently taken off the list. None of these pilots participated in the experiment. Traffic alert and collision avoidance system (TCAS) items

were also removed, because at the time of this study it had not been fully implemented on all commercial aircraft. Elimination of the above items reduced the total set of information elements to 259. The items that remained on the list were considered to be generic. That is, information was currently available and understood by any commercial pilot flying, regardless of aircraft type.

Two stimulus sets, each containing 20 items, were constructed by random item selection from the generic list. Random item selection was used so as not to bias the outcome of the results. The experimental analyses employed herein are based on determining the subject's cognitive organization of the stimuli. Had the experimenters selected the stimuli, it could be argued that the findings were their mental organization of the information.

Although it is desirable to have more items in the stimulus set to enhance the diversity of stimuli, pretesting revealed that more than 20 items took a considerable amount of time, which led to subject fatigue. Twenty items represented a balance between having sufficient stimuli for analysis, and minimizing subject fatigue and time required to perform the tasks. The final sets are hereafter referred to as set A and set B. Four terms were common to both sets and are presented at the bottom of each list in boldface. The two sets are shown in table 1.

Procedure

Each pilot participated in two separate tasks with either set A or set B. The first task required subjects to make similarity judgments between pairs of information elements for a given set. The second task required subjects to rank the terms in their perceived order of priority under two different contexts. Each task is described in detail in the following paragraphs.

For the similarity judgment task, two information elements were presented on a computer screen, one above the other. The subject was then asked to provide a rating of how similar the items were thought to be. A scale from 1 (very similar) to 9 (very dissimilar) was used for the rating. In the lower portion of the display, a scale bar appeared on every trial, which afforded the subjects constant access to the rating scale. Subjects could make similarity ratings by using the number keys or moving an "X" along the rating scale bar with the arrow keys on the computer keyboard and then pressing the enter key once the cursor was at the selected similarity rating. This screen arrangement is shown in figure 1. A computer program (Ricks 1994) randomized the order of the pairs presented to each subject. Subjects rated the similarity of each pair once. This produced a total of 190 trials (half the matrix minus the diagonal). Subjects were free to take a break at any point during the experimental trials when they became fatigued.

To acquaint themselves with the computer display and response arrangement, subjects initially practiced making comparisons among six common automobiles. At the end of the practice, subjects were shown a list of six information elements similar to those to be evaluated in the experimental set. Subjects rated the similarity of these terms, both immediately before and after the

Table 1. Random Samples of Flight-Deck Information Elements

| Set A Flight-Deck Elements	Set B Flight-Deck Elements
Alternate flap control status	Actual EPR
Clock time	Autothrottle disconnect status
Commanded pitch angle	Current bank angle
Crew oxygen flow indication	Current calibrated (indicated) airspeed
Current stab trim	Engine fire condition
Distance to reach selected altitude	Flight path angle (inertial)
Flight number	ILS tuning data (ident/freq/inbd crs/DME)
N1 RPM limit	Incoming ATC message (clearances)
Passenger oxygen system mode (RESET/NORM/ON)	Incoming dispatch message (company comm.)
Predicted fuel at waypoints	IRS information source (for Captn & F.O. PFD/ND)
Relative bearing/distance of waypoints (intersections)	Landing reference speed
Selected roll mode	Oxygen pressure (crew and passenger)
Speed range for in-flight engine restart	Selected vertical speed
Total air temperature	Target N1
True heading (current)	VOR tuning data (freq/ident and radial/course for L/R VOR's)
Zero fuel weight	Wing anti-ice status
Pilot-initiated request or message	**Pilot-initiated request or message**
Predicted/estimated wind for descent	**Predicted/estimated wind for descent**
Selected altitude	**Selected altitude**
Speed restriction data (speed limit, altitude)	**Speed restriction data (speed limit, altitude)**
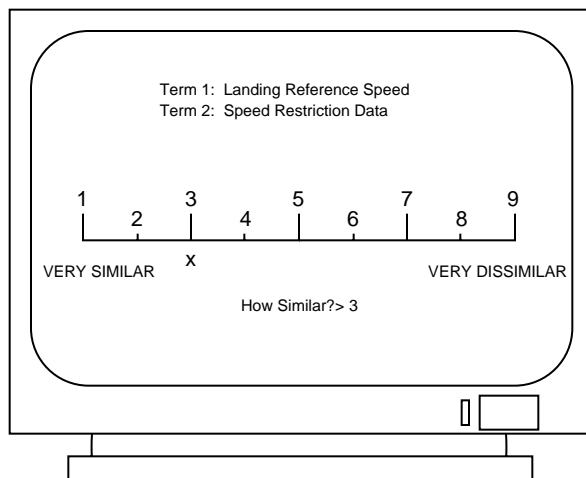
Figure 1. Experimental screen response arrangement.

experimental trials. This practice rating allowed for the calculation of a test-retest correlation to determine whether subjects were using the rating scale consistently.

Prior to making similarity judgments for the experimental set, subjects were given a list of the 20 terms they would be rating. If the subjects had any questions about items on the list, the experimenters showed them the relevant section of the Boeing flight-deck operations manual (Anon. 1992) for clarification.

Subjects were given no specific definition of similarity to perform the comparison task; they were simply told to use whatever definition of similarity they felt comfortable with. Following the experimental trials, however, subjects were asked to explicitly state how they judged the similarity of the presented items. The entire similarity rating task took approximately 30 min.

In the second part of the experiment, subjects prioritized the same set of 20 items that was used for the similarity judgments. These prioritizations were performed under two conditions. Under the first condition, subjects were given the stimuli on 20 separate index cards and told to order the deck of cards according to their perceived priority without regard to a specific phase of flight. This was termed the generic condition. Upon completion of this task, the data were recorded, and the cards were shuffled and given back to the subjects for a second prioritization. Under the second condition, subjects were told to order the items in the context of the takeoff phase of flight. For both conditions, subjects were told to consider all systems as operating normally.

Analogous to the similarity rating task, subjects were told to order the stimuli in each condition by using whatever definition of priority they had. After each ordering, subjects were asked to describe how they prioritized the

data. Subjects were not told of the second prioritization condition (takeoff condition) prior to its introduction. The takeoff condition always followed the generic condition.

Following the experimental tasks, subjects were asked to fill out a questionnaire for their opinion on several issues related to the categorization and prioritization of flight-deck information. A copy of the questionnaire appears in appendix B.

Results and Discussion

Of the 52 subjects retained for analysis, 27 used set A stimuli while 25 used the set B stimuli. The correlations between the sessions preceding and following the experimental trials (six information elements similar to those in the experimental trials) were calculated for each subject. None of the individual correlations were negative, and the average test-retest correlation was a moderate 0.536. Both factors suggest that subjects were using the scale consistently. Given this information, data from all 52 subjects were used.

Upon completion of the experimental trials, data from each subject were stored in matrix format on computer disk. The individual subject files were analyzed with personal computer versions of the Individual Differences Scaling (INDSCAL, Carroll and Chang 1970), hierarchical clustering (average linkage method, cf. Romesburg 1984), and multidimensional preference analysis (cf. Carroll 1972 for analysis implemented with PCPREF) programs. These analyses were run separately on the set A and the set B respondent data, and will be reported separately. Results have been organized in this section by the research question being addressed. Analysis procedures used to address each issue have been broken out separately within each area.

How Pilots Cognitively Process Flight-Deck Information

As noted in the section entitled "Background," cognitive researchers have used psychometric techniques, such as MDS and cluster analysis, to infer how individuals mentally represent information and the associated cognitive processes they use to act upon their representation. In the current study, these techniques were used to examine how pilots represent, and therefore process, typical flight-deck information. Similarity ratings, which were collected by having subjects evaluate the flight-deck information pairs, served as data for both statistical techniques. These data were analyzed by using INDSCAL and hierarchical clustering techniques. For each technique, results will be followed by interpretation and discussion.

Individual Difference Scaling

INDSCAL is a statistical technique that uses proximity matrices as input for each subject. Such matrices represent some measurement of stimulus dissimilarity. In the present case, this measurement was the pairwise comparison data from the computer trials. The INDSCAL model computes a group solution that represents the stimuli spatially in n dimensions. The model attempts to fit the data so that items perceived as being very similar to one another are located in close proximity in the n-dimensional space; whereas, items perceived as being dissimilar lie far apart in the space. Because this solution is generated with data from all subjects, these dimensions are said to be common to all individuals in the sample.

A unique aspect of the INDSCAL approach, however, is the idea that while individuals may process stimuli by using the same group dimensions, it is unlikely that they use the dimensions equally or in the same way. For this reason, INDSCAL also calculates dimensional weights for each subject, which indicate how relevant or salient a dimension is for an individual. Thus, the group space allows the investigator to examine the common cognitive processing of stimuli, while the weights allow one to evaluate the degree to which each subject uses the group dimensions. The subject weights from these INDSCAL analyses are reported in the section entitled "Individual Difference Scaling."

The INDSCAL model was run on the individual subject matrices for set A and set B stimulus sets. Data in each set were scaled in one, two, three, and four dimensions. The dimensional solution was selected based on the percentage of variance for which it accounted and ease of interpretation. Beyond the three-dimensional solution for each set, neither the percentage of variance nor the ease of interpretation increased appreciably; hence, the three-dimensional solutions were retained for analysis. To avoid local minima solutions, multiple runs were performed on each set with five randomly generated starting values. The five random starting values produced virtually identical final configurations within each set, which suggests that none of the final configurations arose from a local minimum (differences among solutions in terms of variance accounted for were in the 1-percent range).

The set A solution selected for interpretation accounted for 36 percent of the variance and the average correlation between computed and actual subject scores was 0.595. The set B solution selected for interpretation accounted for 42 percent of the variance and the average correlation between computed and actual subject scores was 0.635. Given the random stimulus selection, the number of subjects within each set, and the unique nature of some comparisons for the pilots, these fits are good. These three-dimensional solutions, as well as an interpretation for sets A and B, are presented in figures 2 and 3, respectively.

For the plots shown in figures 2 and 3, each stimulus element was plotted in the three-dimensional space and the viewing angle was changed so that each stimulus was visible in the plot. Because this rotation was performed about the coordinate axis, the relative position of items in the plot remained identical. These plots were then shown, without the dimensional labels, to several subject matter experts for interpretation. In addition, the questionnaire data on similarity were used to aid in interpreting these solutions. The names assigned to the dimensions here represent a consensus of the responses obtained.

For sets A and B, one dimension appears to represent traditional flight functions. This is indicated on both plots with items on the left side of the dimension relating to aviate functions, those in the middle relating to navigation, while items on the far right related to communication. This dimension has been labeled "Flight function."

The second dimension appears to represent control and planning functions or, as indicated on the plots, what might more generically be referred to as a tactical and strategic continuum. Note that items on one end are related to direct, short-term, aircraft control (such as "Commanded Pitch Angle" and "Selected Roll Mode" for set A and "Engine Fire Condition" and "Current Airspeed" for set B), while items on the opposite end of this dimension support long-term flight activities (such as "Relative Bearing/Distance of Waypoints" and "Predicted Fuel at Waypoints" for set A and "VOR Tuning Data" and "Predicted/Estimated Wind for Descent" for set B). This latter dimension has been labeled "Flight action."

The third dimension for sets A and B might best be described as the number of times the pilot refers to or uses a given piece of information. As indicated on the plots, this dimension has been labeled "Sample Rate." Items with high values on this scale represent information referred to more frequently than information with lower values. In set A, the "Relative Bearing/Distance of Waypoints" has a higher value on this dimension than "Crew Oxygen Flow Information" or "Passenger Oxygen System Mode." Analogously, for set B, "Current Airspeed" and "Selected Vertical Speed" have high values (representing frequent referral), while items less often used, such as "Engine Fire Condition" and "Wing Anti-Ice Status," have low values along this dimension.

The INDSCAL dimensions discussed previously represent how this group of subjects viewed the
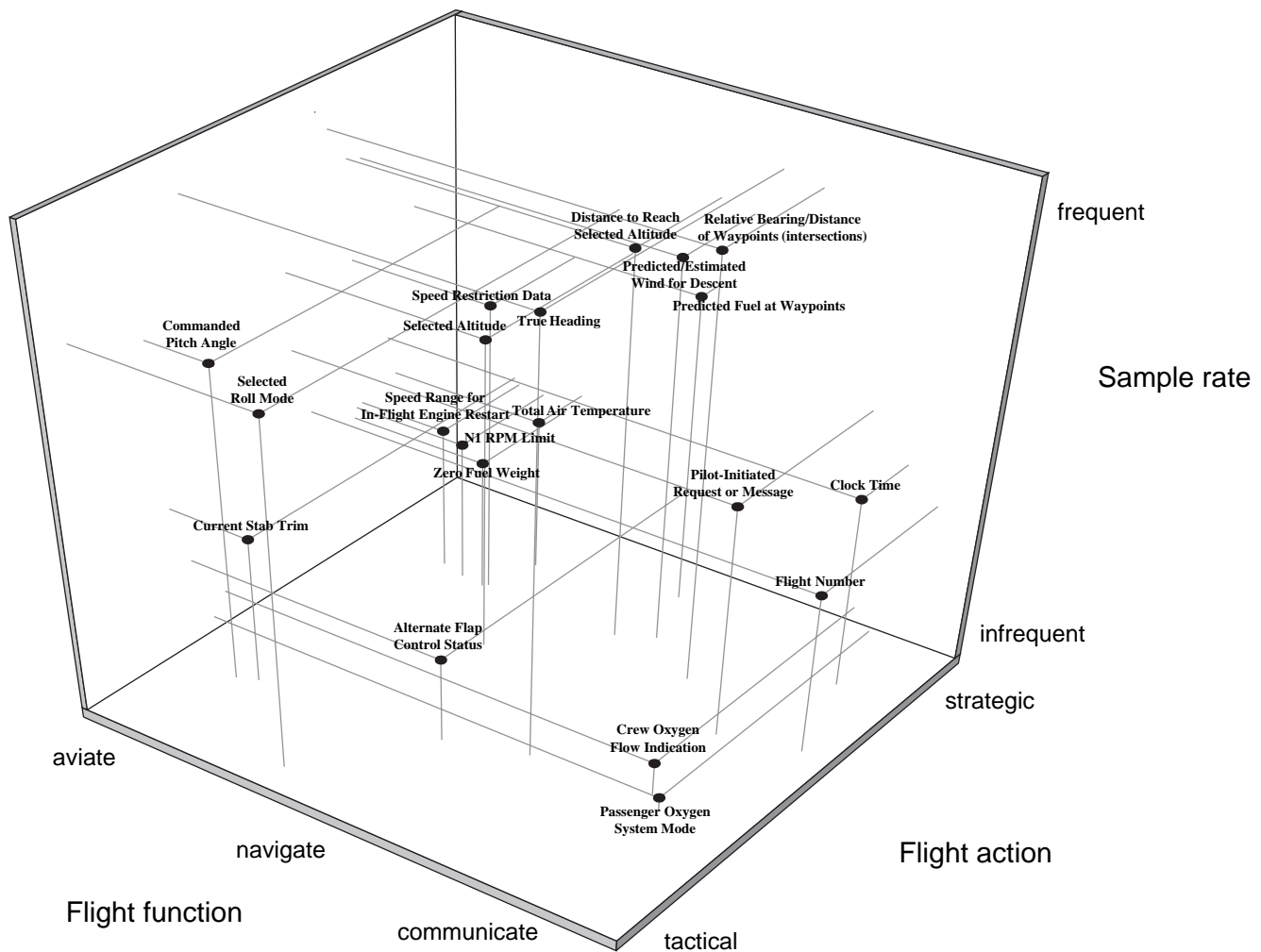
Figure 2. Spatial solution and interpretation of set A similarity data.

information they regularly use on the flight deck. The results for sets A and B are based on a randomly selected sampling of information elements (20 in each set) from the larger set of 259 items. While it is reassuring that the two sets that used these information elements produced similar solutions, the intent is to extend these results to include flight-deck information that was not part of the original set A or B data. An important issue will be how to extend these results to all data.

To illustrate this issue with a concrete example, consider a multidimensional scaling analysis of 10 cities in the United States, where the data matrix consists of distances between those cities. The analysis produces two dimensions; interpretation of which reveals that they correspond to the North/South and East/West map orientations (Kruskal and Wish 1978). Now assume a new city is presented and it is to be placed among the other 10 in the spatial plot. With no knowledge of geography of the

United States and no information about this city in relation to the other cities, this task is not possible.

Note that the INDSCAL dimensional solutions are important, not because they describe the relationship of those items measured in the original analysis; rather, they are important because they give insight into how individuals perceive the relationship of the items and offer possibilities for how this view may be employed in future studies. This view, of course, represents the dimensions. Because the dimensions described here may be representative of cognitive processes that pilots use with flight-deck information, the identification of dimensions is extremely useful. As discussed in appendix A, these dimensions can be quantified and may be employed by having pilots scale other flight-deck information not included in the study as part of the current stimulus set. (While it is possible with the introduction of new technology that completely novel information may be
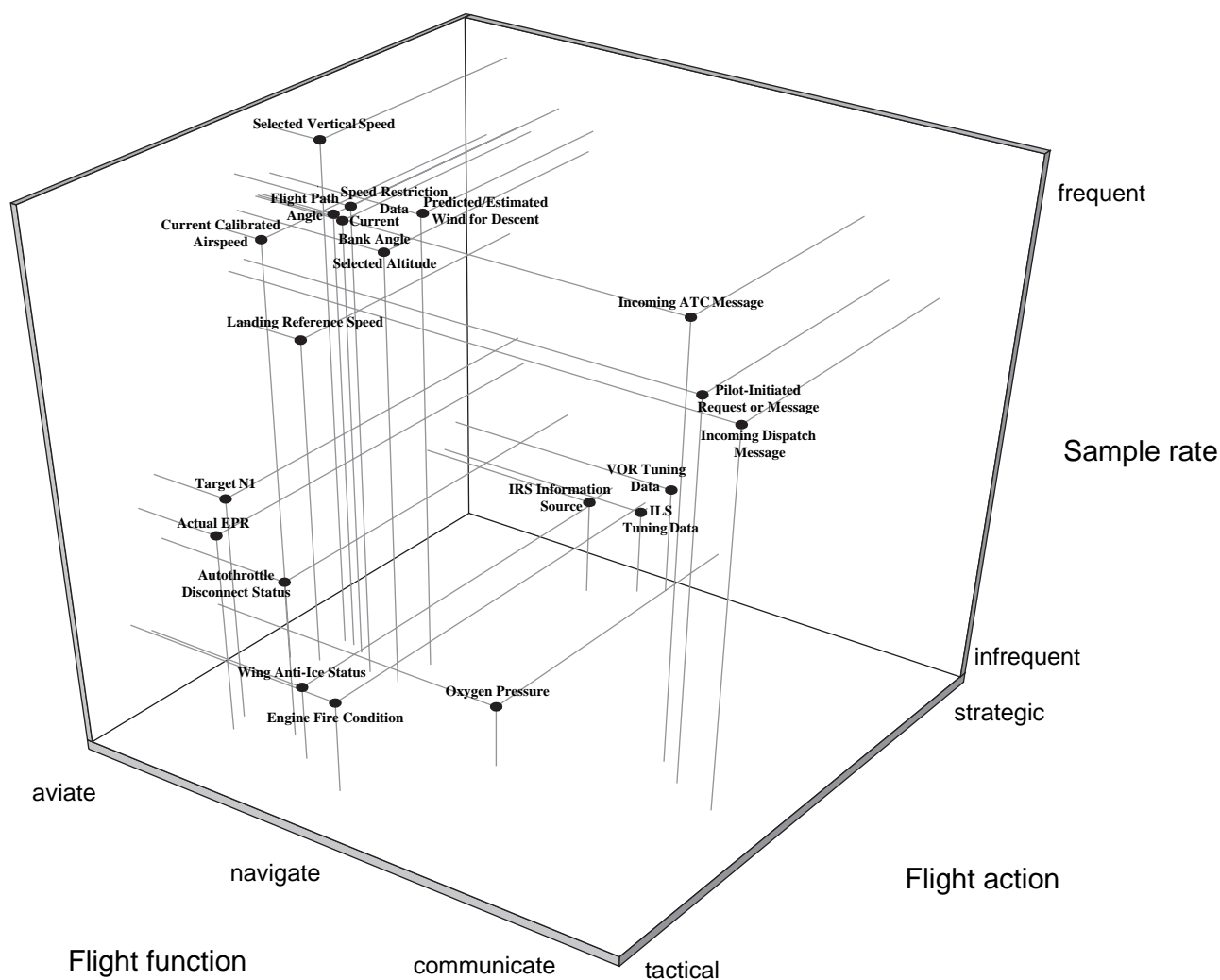
Figure 3. Spatial solution and interpretation of set B similarity data.

introduced, it would seem that a flight-deck engineer could make a fairly accurate estimate of the characteristics of this information.) These data, in turn, could be used as part of a cognitive task analysis for future flight-deck designs. Development and validation of such scales will be, of course, a necessary first step.

Hierarchical Clustering

The MDS analyses provide spatial solutions for the proximity data. When viewing these plots, the dimensions are interpreted primarily by examining where information elements lie along the dimensions. In doing this examination, researchers will often notice items that tend to lie close to one another in the space. Such groups of points, while potentially aiding interpretation of the spatial plots, may also be indicative of a natural category. Such categories emerge from the classifications individuals have learned to make of events or objects encountered in the world. For example, individuals have well-developed categories into which they place animals regularly encountered. These categories could include "cats," "dogs," or "birds." As noted in the "Background" section, this categorizing considerably simplifies an individual's processing of new events and information. Analogously, it seems plausible to assume pilots have high-level categories for flight-deck information they regularly encounter.

This analogy is precisely the reason for exploring those categories into which pilots place flight-deck information. Recall from the "Background" section that studies have shown that presentation of information consistent with an individual's cognitive categorization scheme may lead to superior performance, as indicated by measures for accuracy and response speed. For this reason, identification of the natural categories that are used by pilots (if they indeed exist and can be adequately

defined) may prove useful in developing new information presentation formats and information management systems.

To address this issue, the original (or raw) proximities for the pairwise comparison data, which were averaged across subjects were clustered hierarchically. Analyses were run with the PROC CLUSTER procedure developed by SAS Institute Inc. (Anon. 1988), which specified the average linkage method. Cluster analysis is a technique that attempts to group objects together according to their similarity. Although several variations are possible, the analysis typically begins by grouping the most closely related objects (based on proximity), treating this cluster as a new individual object, and then recalculating its proximities with the remaining objects. This iterative approach continues until only two clusters remain. (For a full description of the most widely used clustering approaches, see Romesburg (1984).) The resulting output is often presented in the form of a dendrogram (or hierarchical tree). When presented laterally, as is the case here, closely related items cluster together at the far right of the dendrogram, while the further one proceeds toward the left (up the plot), the less similarity is perceived among individual items and the fewer distinctions (in the form of clusters) can be made.

Cuts in the dendrogram were made where there was a high level of stability in the obtained clusters. This criterion means that the cut point could be moved across a relatively wide range of the dendrogram without significantly affecting the number of clusters (for information on how to make dendrogram cuts, see Romesburg 1984). The dendrograms for the raw similarity ratings of sets A and B are presented in figures 4 and 5, respectively. Based on where the cut was made, boxes have been placed around the clusters. Interpretations have also been assigned to these clusters and are so identified. As with the INDSCAL plots, the clusters were examined by subject matter experts and names were assigned to the clusters after review.

Although some differences occur between the spatial position of the information elements from the INDSCAL analyses and the clustering results, in general the pattern of results proved quite similar. That is, items that were located close together in the INDSCAL plot also tended to group together in the cluster analyses.

Based on the dendrogram cuts, four high-level clusters emerged for the set A raw proximities data and have been labeled in figure 4. These clusters include

1. Aviation

2. Navigation

3. Communication

4. Systems administration

Each of these categories may be defined by using a modified series of definitions described in Abbott (1993). Aviation may be defined as the process "of adjusting or maintaining the flight path, attitude, and speed of the [aircraft] relative to flight guidance requirements." Navigation may be defined as the process of "developing the desired plan of flight. . . and monitoring its progress." Communications can be defined as involving the transfer of information between the crew and ATC, the airline company, and flight crew members. Finally, systems administration may be defined as the process of monitoring the state of aircraft systems, identifying when actions may need to be taken to return a system to nominal status (if possible) and making such changes.

Closer inspection of the large navigational cluster suggests a further distinction could be made within that grouping. A differentiation could be made between pure navigational information ("Relative Bearing/Distance of Waypoints" and "Distance to Selected Altitude") and information traditionally considered as being reference data ("Speed Range for In-Flight Engine Restart" and "N1 RPM Limit"). While it is useful to recognize that such a distinction can be made, for purposes of parsimony, the four-cluster solution was retained. Based on the dendrogram cuts for the raw proximities data, the same four high-level clusters emerged in set B as emerged in set A.

Results from the cluster analyses suggest pilots conceptualize flight-deck information into at least four distinct categories. These categories are probably best captured by the descriptors (1) aviation, (2) navigation, (3) communication, and (4) systems administration. Relating these results to the prior INDSCAL solutions, these categories appear most strongly related to the flight function dimension. While three distinct activities appeared along that dimension for the INDSCAL analyses, the cluster results suggest that systems administration would most closely align itself with communication. Finally, the fact that similar clusters emerged across two separate data sets, with each having substantially different items, is reassuring because it suggests the same clusters generalize across two data sets.

While three of the four information elements common to both sets appeared in different clusters, it is important to note this occurrence was most likely because of the random stimulus selection employed. For example, a cluster, most appropriately considered as navigational, emerges in both sets A and B. However,
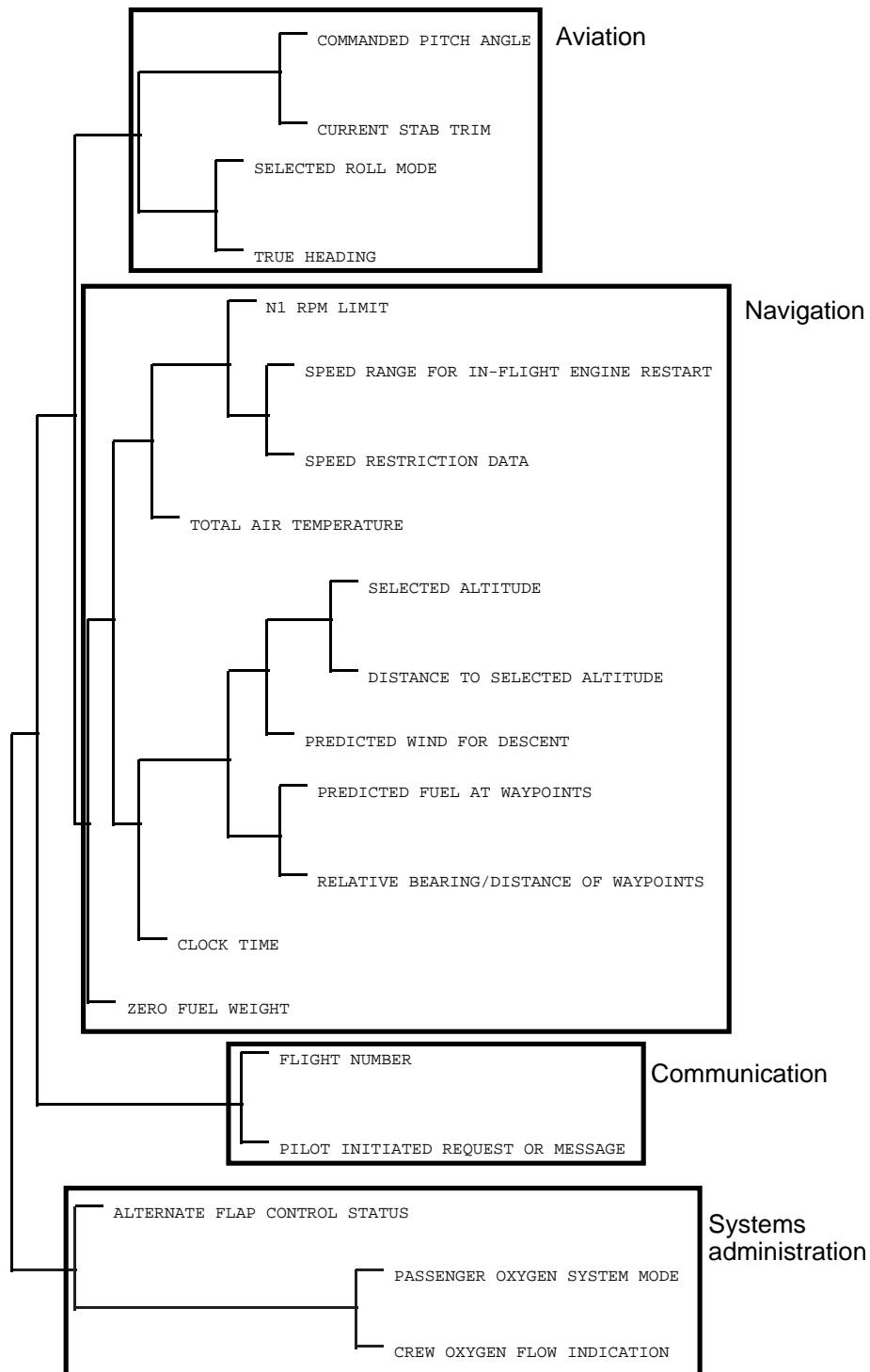
Aviation

COMMANDED PITCH ANGLE

CURRENT STAB TRIM

SELECTED ROLL MODE

TRUE HEADING

Navigation

N1 RPM LIMIT

SPEED RANGE FOR IN-FLIGHT ENGINE RESTART

SPEED RESTRICTION DATA

TOTAL AIR TEMPERATURE

SELECTED ALTITUDE

DISTANCE TO SELECTED ALTITUDE

PREDICTED WIND FOR DESCENT

PREDICTED FUEL AT WAYPOINTS

RELATIVE BEARING/DISTANCE OF WAYPOINTS

CLOCK TIME

ZERO FUEL WEIGHT

Communication

FLIGHT NUMBER

PILOT INITIATED REQUEST OR MESSAGE

Systems
administration

ALTERNATE FLAP CONTROL STATUS

PASSENGER OXYGEN SYSTEM MODE

CREW OXYGEN FLOW INDICATION

Figure 4. Clustering dendrogram of set A similarity ratings.

CURRENT BANK ANGLE

FLIGHT PATH ANGLE

CURRENT CALIBRATED AIRSPEED

SPEED RESTRICTION DATA

LANDING REFERENCE SPEED

SELECTED ALTITUDE

SELECTED VERTICAL SPEED

PREDICTED WIND FOR DESCENT

TARGET N1

ACTUAL EPR

AUTOTHROTTLE DISCONNECT STATUS

ENGINE FIRE CONDITION

WING ANTI-ICE STATUS

Aviation

VOR TUNING DATA

ILS TUNING DATA

IRS INFORMATION SOURCE

Navigation

PILOT INITIATED REQUEST OR MESSAGE

INCOMING ATC MESSAGES

INCOMING DISPATCH MESSAGE

Communication

OXYGEN PRESSURE

Systems
administration

Figure 5. Clustering dendrogram of set B similarity ratings.

because the stimulus selection procedure selects randomly, set B had three information elements, that were all strongly related to flight navigation. These relationships had the effect of forming an extremely tight navigation cluster that contained three navigation items and forced the other items, that were related to navigation, albeit less so, into the larger aviation cluster. Likewise in set A, four information elements, strongly related to aviation, formed a tight cluster that left a large cluster that contained items principally related to navigation. This failure to find the identical information elements in the same clusters is not surprising because the sets were randomly formed, and the experimenters had no control over the combination of stimuli appearing in each set. When examining clustering results, the important thing to note is that the cluster names are consistent with the items they comprise.

The cluster analyses also provide insight into how pilots categorize regularly used flight-deck information. While most anecdotal evidence from pilots suggests that they divide information into three categories (aviate, navigate, and communicate), the current findings suggest that a richer categorization scheme may be operating, with pilots willing (either consciously or not) to make finer gradations to the information they process. These gradations are seen here as systems administration emerging as a fourth category in the analyses and within categories, such as navigation, the emergence of further differentiation in the form of navigational information and reference information. (An alternative distinction within this category would be dynamic and static information, because items such as "N1 RPM Limit" and "Speed Restriction Data" involve fixed values, while "Predicted Wind for Descent," and "Predicted Fuel at Waypoints" involve dynamic or changing information.)

In determining how pilots process (and prioritize) information, it would be tremendously useful to work with high-level functionally related categories, such as those empirically found herein instead of working with the individual information elements. As will be discussed in the section "How Pilots Prioritize Flight-Deck Information," the clusters emerging from these analyses may prove useful in developing models for the prioritization of flight-deck information.

How Pilots Prioritize Flight-Deck Information

How pilots prioritize flight-deck information was a second concern of this research project. While the procedures already described (INDSCAL and hierarchical clustering) provide the researcher with insight into how pilots perceive the flight-deck information elements to be related, these procedures do not directly address the underlying method that is used to prioritize the informa-

tion. To explore this, a statistical technique called multidimensional preference analysis was employed. As noted earlier, to address the dynamic nature of flight-deck information usage, pilots were asked to prioritize under two separate conditions. One prioritization occurred without regard to any specific phase of flight (generic condition), while the other was in the context of the takeoff condition of a flight mission. The results for each of these experimental conditions will be presented separately for sets A and B.

Multidimensional Preference Analysis (MDPREF) is a technique that takes ranked data and provides a joint representation of the data and the subject preferences. The analysis presents the stimuli spatially, analogous to multidimensional scaling. The MDPREF analysis was conducted with the PCPREF algorithm, which is a personal computer version of MDPREF. (See Carroll 1972). This implementation accepts data ordered by subjects as input. The program then computes principal components from these data. The number of components that are retained is determined by (1) examining the additional variance that is explained by adding a new component and (2) the overall interpretability of the solution.

MDPREF analyses were conducted on each set for both generic and takeoff conditions. Across all four runs, the dimensionalities were extremely similar. Each solution generated one principal component that accounted for over 40 percent of the variance and a second component that accounted for approximately 10 percent of the variance. These variance percentages indicate the extent to which each component provides information about the data. When the components are viewed in the context of uncovering information in the raw data matrix, large percentage variances could be expected to account for more information. As successive components are extracted, the variance accounted for by each component will decline, and begin to represent random variation. Here, the first components accounted for a large amount of the variance for each solution, and the variance of components beyond the first two dropped off sharply. For these reasons, and to achieve maximum interpretability, two-dimensional solutions were selected for all runs. The percentage variance accounted for by each factor is presented in table 2 for sets A and B for the generic and the takeoff conditions.

An overview of the results for these four factorial conditions will be presented because the same pattern of results emerges across all four conditions. Specific results for each of the four conditions will be discussed in further detail in the following sections. Along the first dimension, high values were found for aviation information, moderate values were observed for navigational information, and system status information exhibited the

Condition	Set A		Set B	
	Dimension 1	Dimension 2	Dimension 1	Dimension 2
Generic	0.434	0.138	0.478	0.094
Takeoff	0.504	0.130	0.595	0.109

lowest values along this dimension. This latter finding is not surprising, given that status data in these sets dealt primarily with emergency information and subjects were instructed under both conditions (generic and takeoff) to consider all systems as operating normally. The second dimension appears to be capturing communications information because these items define the extreme end of the dimension.

Multidimensional Preference Analysis of Set A Rankings

Average rank of set A data under both the generic and the takeoff conditions are provided in table 3.

Table 3. Average Rank of Set A Data for Generic and Takeoff Conditions

Set A data	Generic condition	Takeoff condition
Alt	1	4
ComPitch	2	1
TrueHead	3	5
RollMode	4	7
N1 RPM Limit	5	2
SpdRestriction	6	11
RelBearing	7	14
PredFuel	8	17
CurrStabTrim	9	3
DistToAlt	10	12
ZeroFuelWt	11	6
Plt Reqt/Mess	12	9
Flt #	13	8
CrewOxygen	14	16
PredWind	15	20
TAT	16	10
SpdRange	17	18
PassOxygen	18	15
Clock	19	13
AltFlapStatus	20	19

The correlation between the generic and the takeoff rankings was 0.701, which indicates that the same information judged to be important during takeoff is also generally seen as important in the generic phase of flight. To more fully investigate how the information elements were judged for importance, the ranked data were submitted to an MDPREF program for a spatial analysis. These results are given for each condition in the following sections.

Generic condition. Two principal components accounted for a total of 57 percent of the variance. The spatial plot is shown in figure 6. To aid with interpretation of the plot, those clusters obtained from the similarity data for set A (reported in the "Hierarchical Clustering" section) have been labeled with the high-level descriptors (aviation, navigation, communication, and systems administration) used previously. This technique is used because the cluster names superimposed on the plots were derived from a completely independent analysis. Interpretation of spatial solutions is achieved by defining dimensional endpoints and by identifying those items that tend to group together. The cluster analysis provides this information and allows any structure in the solution to be observed.

Items in the high end of the first dimension principally include aviation information. Navigational information falls toward the middle of this dimension, while systems administration information is found in the lower end. The second dimension has communications information at the high end, while systems administration data account for some items at the lower end.

While the dimensions from MDPREF plots often lead to obvious interpretation, the solutions obtained here are more complex. One reasonable interpretation, which appears across all the MDPREF runs, would be a primary importance dimension and a secondary importance dimension. Several factors argue for this interpretation. First, the primary dimension (*X*-axis) captures a large amount of the variance. This occurrence could be taken as suggesting a unidimensional solution is sufficient to capture the information contained in the ordering data. Because the first dimension in an MDPREF solution typically approximates a consensus of the agreement of the stimulus ranking across subjects, this solution would suggest that the first dimension is a good representation of what these pilots consider to be the most important information for safe flight.

Second, if the first dimension corresponds to the pilot's perception of overall information importance, then one could expect to find a high correlation between the average rank of items in the generic condition (from table 3) and their coordinate values along the first dimension. This was in fact the case, with a correlation of −0.94. Because of this high correlation and because
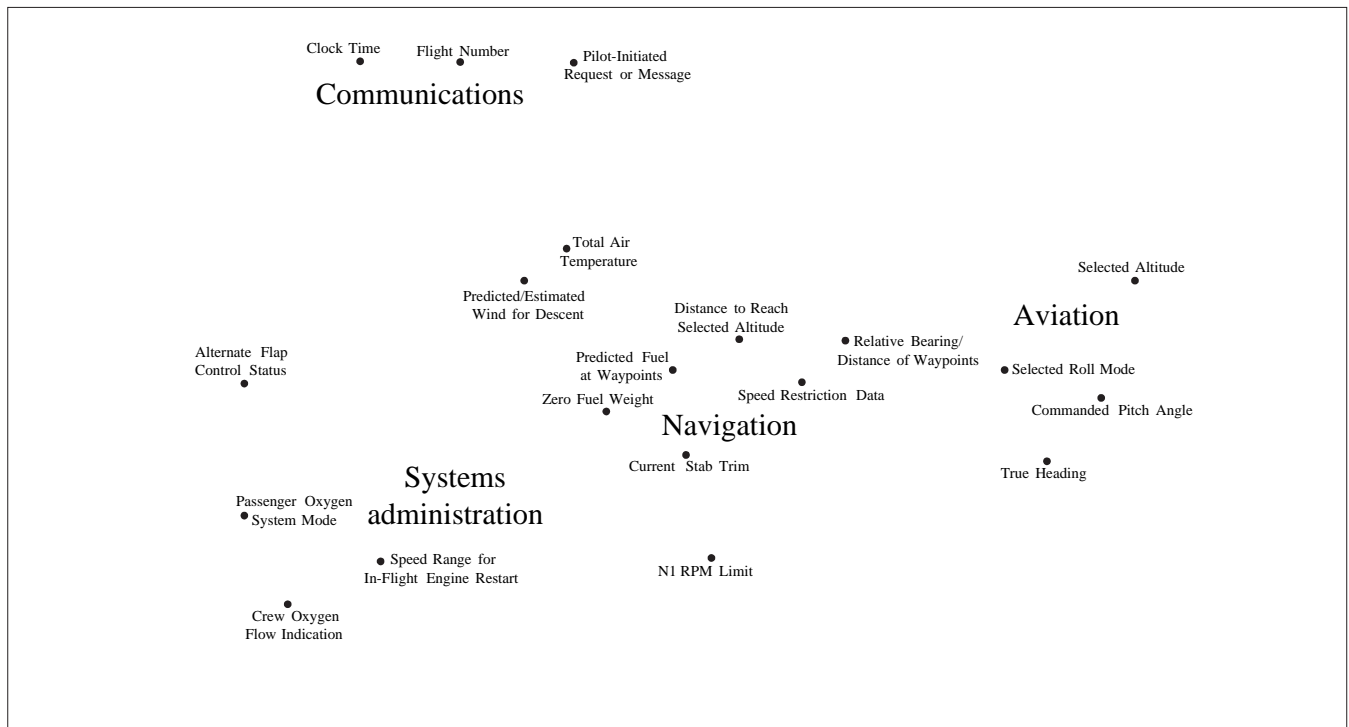
Figure 6. Spatial solution and interpretation of set A ranking data for generic condition.

subjects were told to rank the information in terms of perceived priority, the first dimension (*X*-axis) appears to be measuring an overall information importance factor.

As noted, while the use of previously defined clusters represents a useful method for interpreting MDPREF spatial plots, it is also informative to analyze those clusters that emerge with the MDPREF stimulus coordinates as input data for the analysis. This analysis determines whether similar clusters are derived from entirely different experimental techniques. If so, the relative prioritization of classes of flight-deck information as revealed in the clusters can be examined. In the present case, three distinct clusters emerged that were essentially the same as those found in the set A cluster analyses of the similarity proximity data. The primary difference is that one large cluster emerged that encompassed the items previously appearing under the aviation and navigation clusters.

Takeoff condition. Two principal components accounted for 63 percent of the variance. The spatial plot is shown in figure 7. As in the generic condition, items showing high to moderate values along the first dimension are again included in the aviation category. Navigation elements once again lie in the middle, while systems administration items define the lower end. As before, communications items have large values on the second dimension. Note, however, that in comparing figure 6

with figure 7, certain elements have shifted because of their particular importance for takeoff. This shift is most easily seen with the increasing importance of "Zero Fuel Weight," "N1 RPM Limit," and "Current Stab Trim." Likewise, note that information relevant for in-flight actions moves down in the first dimension (e.g., "Predicted Fuel at Waypoints" and "Predicted/Estimated Wind for Descent"). Finally, the full dimensionality of the data can be exploited by noting, in a two-dimensional solution, information with high values on both dimensions will appear in the upper right hand quadrant of the plot. In this case, those items particularly relevant for takeoff shifted into the upper right quadrant of figure 7. This shift is as expected, given the logic behind the MDPREF analysis.

As in the generic condition, a cluster analysis was performed on the coordinate data obtained from the MDPREF analysis for the takeoff condition. This analysis produced a somewhat different result from that obtained in the generic condition. While the same number of clusters emerged, items appearing within the clusters were more diffuse. Examination of these clusters revealed that they might represent a straightforward prioritization of information required for safe takeoff (analogous to a checklist) with one group containing the most important information, another group containing secondary information, and a third group containing the least important information.
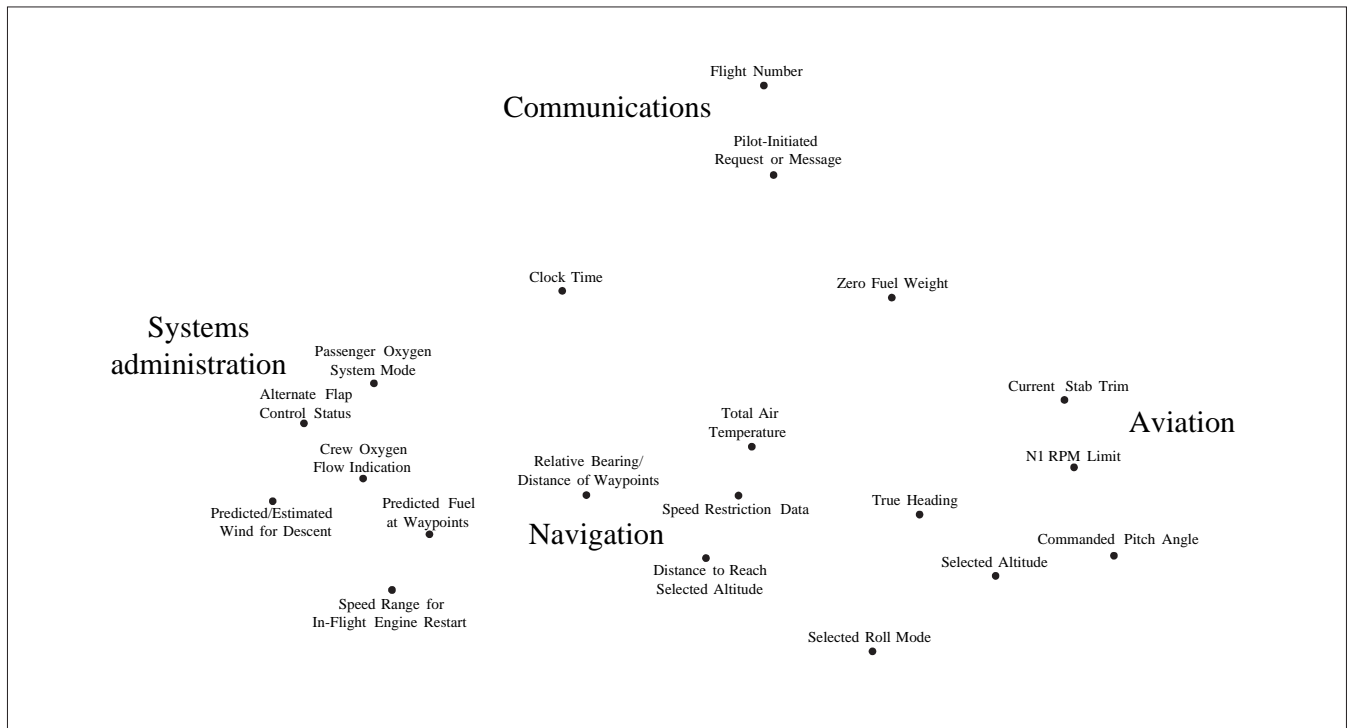
Figure 7. Spatial solution and interpretation of set A ranking data for takeoff condition.

Multidimensional Preference Analysis of Set B Rankings

The average rank of set B data are presented in table 4.

Table 4. Average Rank of Set B Data for Generic and Takeoff Conditions

Set B data	Generic condition	Takeoff condition
Airspeed	1	2
BankAng	2	5
Alt	3	6.5
Actual EPR	4	3
Flight Path Angle	5	8
Target N1	6	1
EngFireCon	7	6.5
SpdRestriction	8	10
Sel Vertical Spd	9	9
LndRefSpd	10	17
Inc ATC Msgs	11	11
ILS Tuning Data	12	16
AThrottle Data	13	4
VOR Tuning Data	14	14
Plt Reqt/Msg	15	12
IRS Info Source	16	15
PredWind	17	19
Anti-Ice	18	13
OxyPressure	19	20
Inc Dispatch Msg	20	18

There was relatively good agreement on the importance of this information, which is indicated by a correlation of 0.811 between the two rankings. This correlation indicates that information deemed to be important during takeoff is also seen as being important during the generic phase of flight. As with the set A data, the subject rankings were analyzed with MDPREF to more fully explore the data structure. These results will be presented separately for each of the two conditions in the following sections.

Generic condition. Two components accounted for 57 percent of the variance. The spatial plot is shown in figure 8. To aid with interpretation of the plot, high-level descriptors for clusters obtained from the similarities proximity data for set B (reported earlier in "Hierarchical Clustering" section) are identified on the spatial plot. Note that the aviation cluster captures the majority of information elements with high values on the first dimension. Navigation and communications data show lower values on this dimension, while the one systems administration item ("Oxygen Pressure") shows an extremely low value. The second dimension is defined almost exclusively at the high end by "Engine Fire Condition." Navigation elements occupy the lower end of this second dimension. As with the set A data, the two dimensions would again appear to represent a distinction between information of primary and secondary importance. The
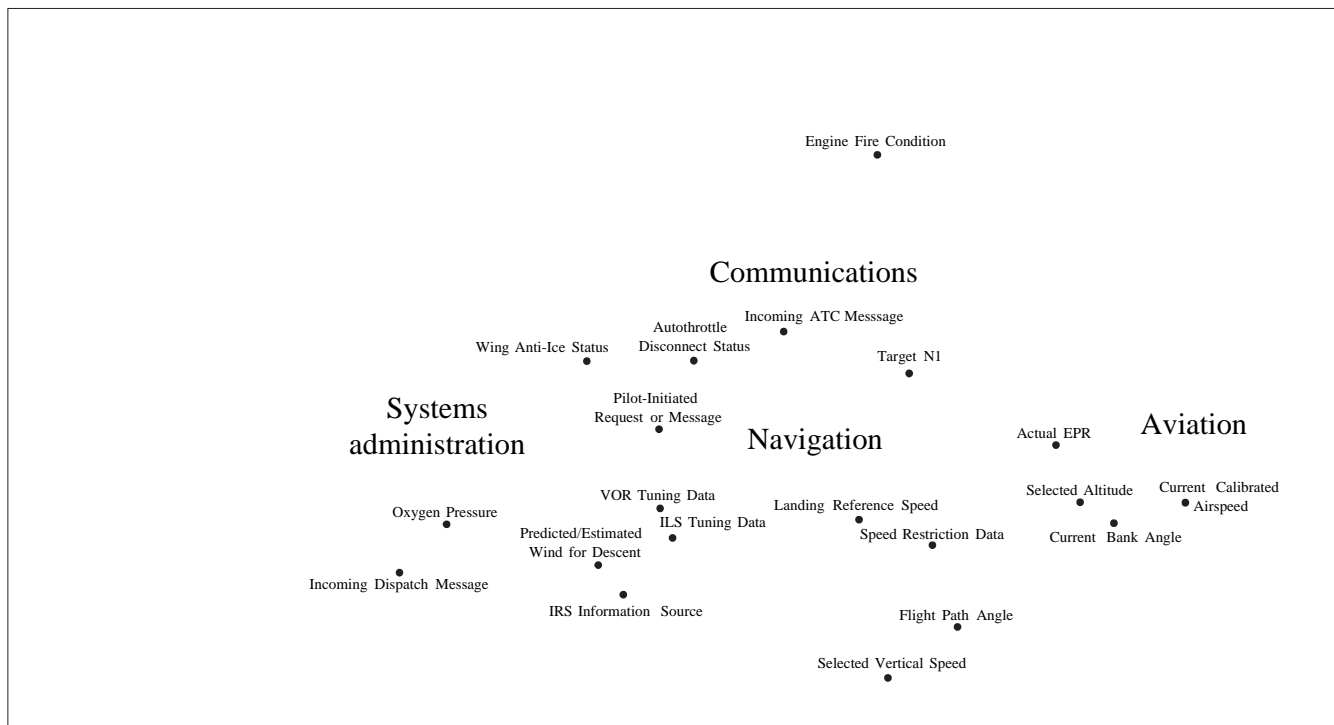
Figure 8. Spatial solution and interpretation of Set B ranking data for generic condition.

correlation between the average rank in the generic condition for each information element and their coordinate values along the first dimension was −0.98.

To more fully analyze these data, a cluster analysis was run on the stimulus coordinates from the MDPREF analysis. Four distinct clusters emerged. Interestingly, this cluster analysis proved to be different from that found for the raw similarity data. (See fig. 5.) Here, the clusters were less well defined, and in some cases, contained different information. These clusters may have resulted from the nature of the primary task. Again, the ranking may become synonymous with relative priority.

Takeoff condition. Two principal components accounted for 70 percent of the variance. The spatial plot is presented in figure 9. The clusters derived from the set B raw proximities have again been identified on the plot. As noted in the generic condition, the aviation cluster contains the majority of items in the stimulus set and again exhibits high values on this first dimension. As before, navigational items lie along the middle of the dimension. The second dimension is defined at the high end by communications information, while navigational data occupies the other end.

This particular data set produced interesting results across the two contextual conditions. First, note that for

the analysis of the takeoff condition with MDPREF, within the aviation cluster, several information elements shifted in perceived importance. Most notably, propulsion information ("Target N1," "Engine Fire Condition," and "Actual EPR") moved toward the high end of the first dimension. Additionally, communication information moved to the top of the second dimension.

To more fully analyze these data, a cluster analysis was run on the stimulus coordinates from the MDPREF analysis. Four distinct clusters again emerged. As in the generic condition for this data set, the clusters obtained here contain some analogous items leading to less well defined categories. As before, this may be due to the particular set of stimulus items and the task requirements.

Discussion

In the MDPREF framework, the content of the first dimension is that which is most important to the subjects; while the second dimension represents the next most important content that is not correlated with the first. To name these dimensions on the spatial plot, an analysis is normally made of those items on the extremes of each dimension of any groups of items clustering together that share obvious similarities. Although a spatial plot is generated, there is no guarantee that a meaningful interpretation will emerge for the dimensions.
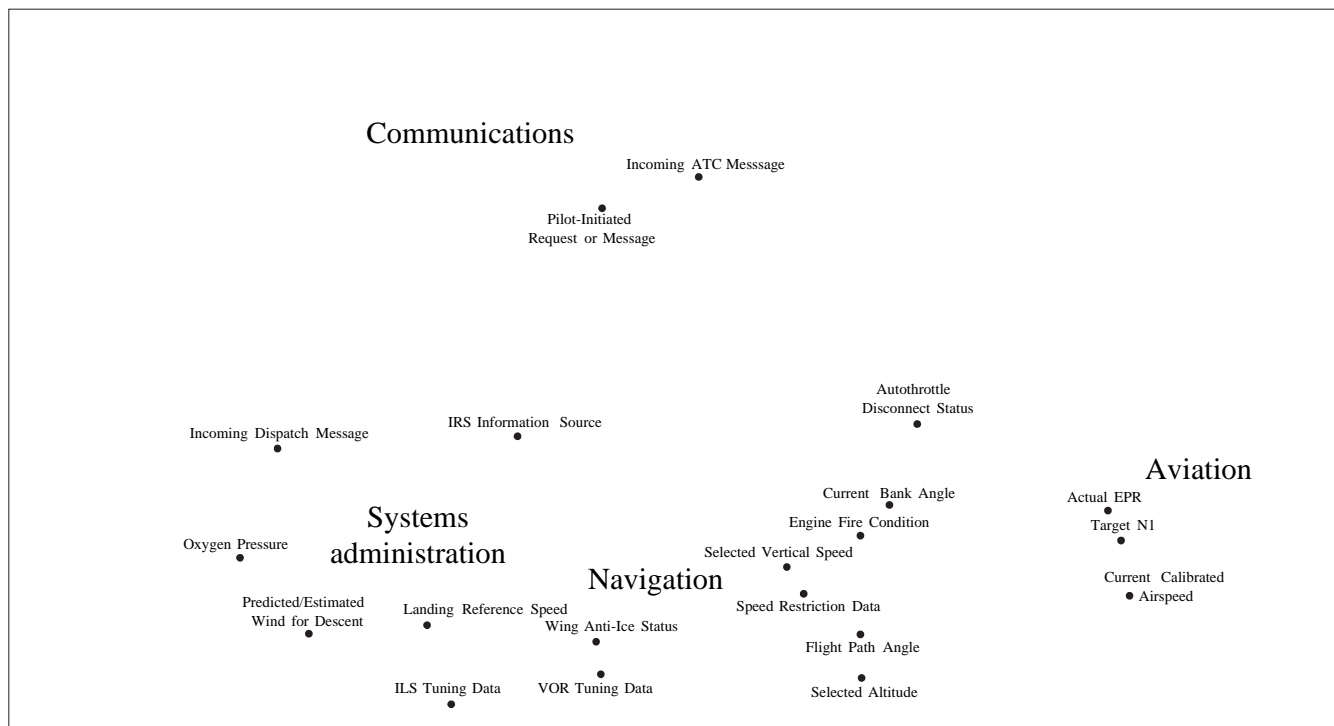
Figure 9. Spatial solution and interpretation of Set B ranking data for takeoff condition.

For both sets, functionally related flight-deck information appears to be collocated in the MDPREF plots. This collocation is independently confirmed by the clusters derived from the similarity data. Information elements in close proximity to one another on these spatial plots (figs. 6 to 9) generally fell within the same four clusters (aviation, navigation, communication, and systems administration) described earlier. These cluster names have been placed in the spatial plots to indicate, generally, where the individual elements composing them are located.

As discussed above, the first, or consensus dimension (*X*-axis) in these analyses, could simply be labeled as a primary importance dimension. The second dimension might represent information considered to be of secondary importance. If interpreted as such, the relative position of items forming clusters directly indicates the information pilots consider to be most critical for safe flight. This relative position of clusters in the MDPREF plots is, perhaps, one of the more important aspects of this analysis. Just as the information elements can be ranked along the MDPREF prioritization dimensions, so too can clusters of information, if the items within each cluster share some logical similarity.

Several additional issues emerge from these analyses. First, it is important to note that the relative lack of importance assigned to the emergency information ele-

ments for the systems administration cluster is almost certainly due to the experimental instructions. Recall that subjects were told to consider all systems as operating normally. In addition to this, those emergency items appearing in sets A and B would not seem particularly relevant for a takeoff emergency situation. Second, note that, for both sets, the first dimension in the context dependent condition accounts for almost as much variance by itself as the two dimensions in the context independent (generic) condition. One explanation for this overwhelming difference in variance is that by providing a specific context in which to prioritize, the experimenter implicitly requires that the pilot give a ranked checklist of that information necessary for a safe takeoff. This ranking can most easily be captured in a single dimension (from most to least important).

Cognitive Processing of Information Among Pilots

Of interest in the study presented herein was the extent to which the cognitive processing of stimuli was similar among pilots. This similarity can be evaluated for both the cognitive representation pilots have for the information elements (as analyzed with the INDSCAL methodology) and the prioritization of flight-deck information (as analyzed with the MDPREF methodology). Results addressing each of these areas will be presented separately.
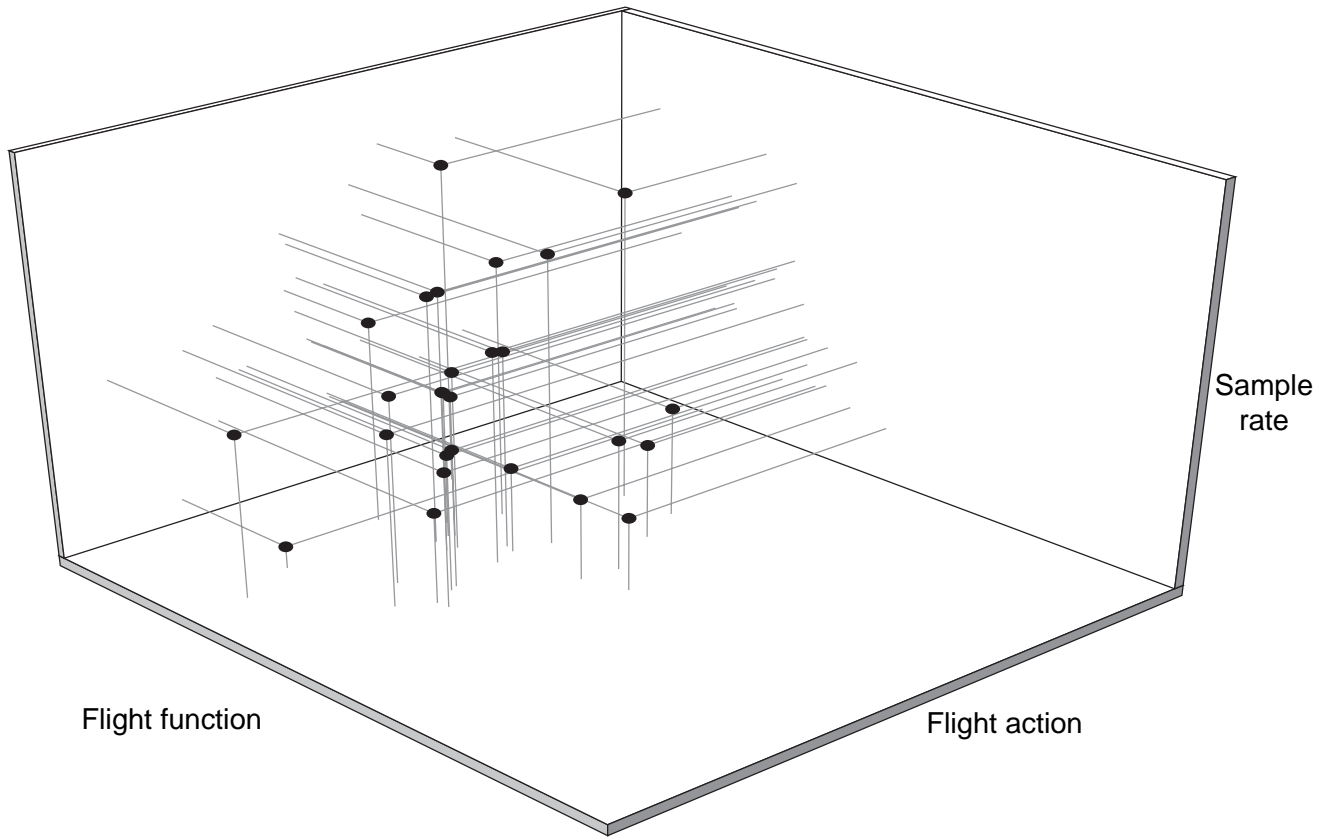
Figure 10. Subject plot for set A similarity ratings.

Individual Difference Scaling

For the similarity judgments, INDSCAL provides a subject weight space that indicates how salient each of the dimensions extracted from the analysis is to an individual subject. For dimensions interpretable as reflecting cognitive processes, salience can be seen as equivalent to the use of the dimensions. If a subject is using all dimensions equally, then the weights given to each dimension should be about equal. If, however, one (or more) of the dimensions is being used more than the others, there should be a difference among the weights. Three-dimensional plots for the sets A and B subject weight spaces are shown in figures 10 and 11, respectively.

The subject weights indicate that for set A, pilots were using all dimensions about equally. For set B, they tended to use dimensions 2 and 3 (flight action and sample rate) less than dimension 1 (flight function). These results were confirmed statistically by analyzing the INDSCAL subject weights in a one-way ANOVA, with each of the three dimensions representing a treatment level. For set A, there was no statistically significant difference among the three dimensions $(F(2, 52) = 1.73, p > 0.05)$. However, for set B, there was a statistically

significant effect $(F(2, 48) = 15.07, p < 0.01)$. In this case, subjects showed higher weights on the first dimension (flight function) than on either of the other two dimensions. These higher weights were confirmed statistically by using Bonferroni post hoc comparisons (see Cliff 1987). While dimension 1 was significantly different from dimensions 2 and 3 $(F(1, 24) = 19.28, p < 0.01$ and $F(1, 24) = 22.82, p < 0.01$, respectively), no statistically significant difference was found between dimensions 2 and 3 $(F(1, 24) = 4.33, p > 0.05)$. Average subject weights for each dimension and associated standard deviations are shown in table 5.

Multidimensional Preference Analyses

In an MDPREF analysis, preference of each subject for the stimuli is indicated by a vector in the spatial plot. Projecting the stimulus points onto the vector would represent the relative preference for the stimuli. Furthermore, the cosine of the angle subtended between a vector and a dimension will provide the correlation of a subject's preferences with that particular dimension. If the spatial dimensions for the group data are interpretable, as they frequently are, then differences among subjects in
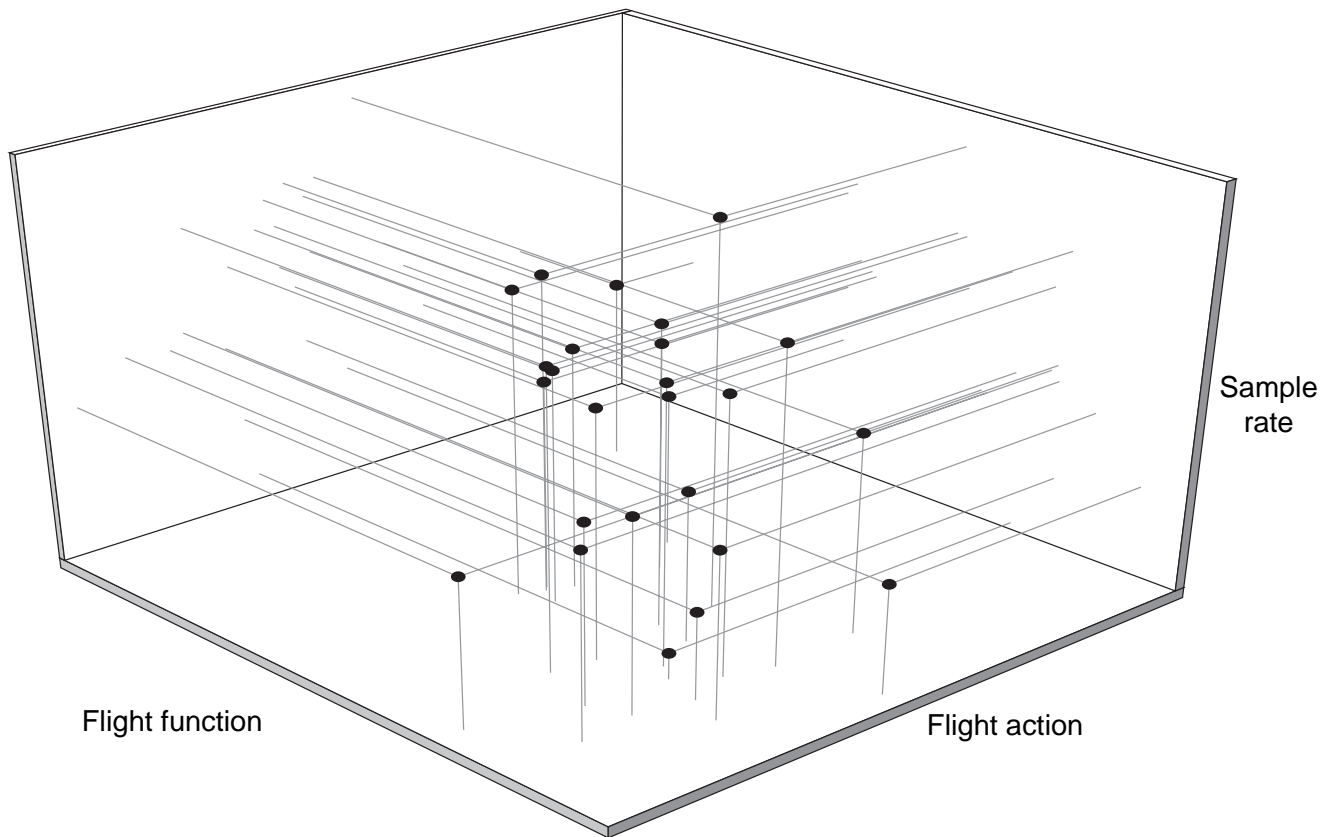
Figure 11. Subject plot for set B similarity ratings.

Table 5. Averages and Standard Deviation of Subject Weights for Sets A and B

	Dimension 1		Dimension 2		Dimension 3	
Set	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
A	0.295	0.075	0.333	0.107	0.354	0.122
B	0.433	0.121	0.338	0.094	0.271	0.107

their use of dimensions can be inferred by the location of subject vectors within the space.

For the MDPREF information prioritization analysis, substantial agreement existed among the subjects. For each of the four solutions considered here, over 70 percent of the subject vectors fell between 45° above and below the first dimension. This position indicates a relatively high degree of agreement among subject responses. Figure 12 plots the set A data for the generic and takeoff conditions. In these plots, each subject is represented as a point. Figure 13 shows the analogous data for the set B stimuli. One can readily observe that the spread of points is tightly arranged around the first dimension for both experimental conditions. The

abscissa may best be thought of as a consensus dimension, which reflects an average ranking by a subject of information elements.

Discussion

In examining the INDSCAL subject weights, differences between sets A and B in the relative use of dimensions by pilots is not entirely clear. However, it is important to note that because the stimulus sets were constructed by using random selection, the particular combination of items in set B may have led to a situation where one dimension became more relevant, interpretable, or salient to the subjects. The anomalous results of sets A and B may be due to the vagaries of the stimuli
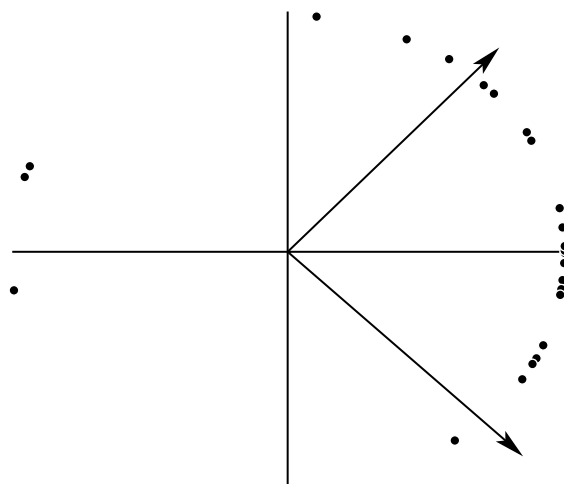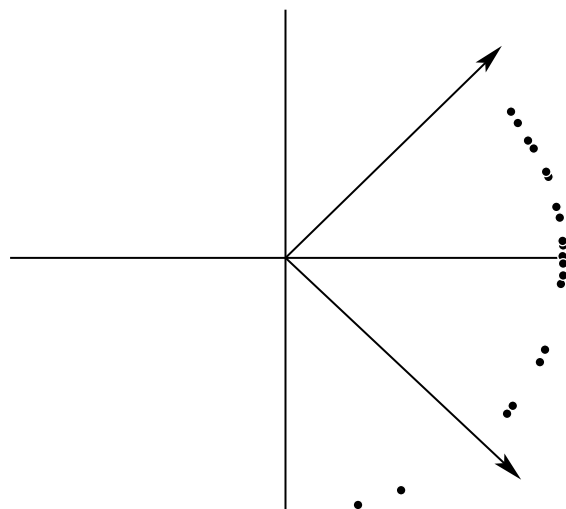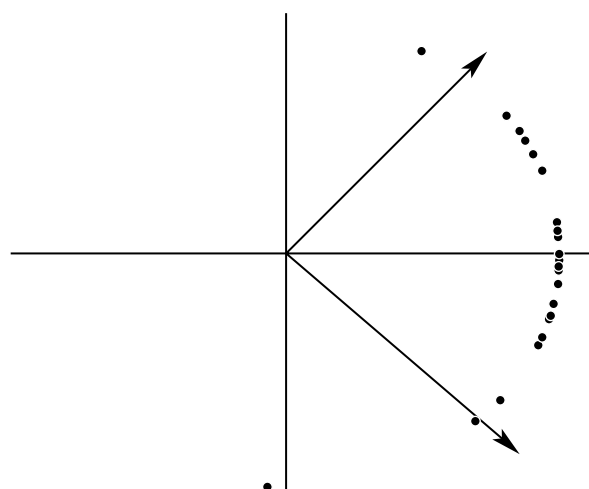
(a) Generic.
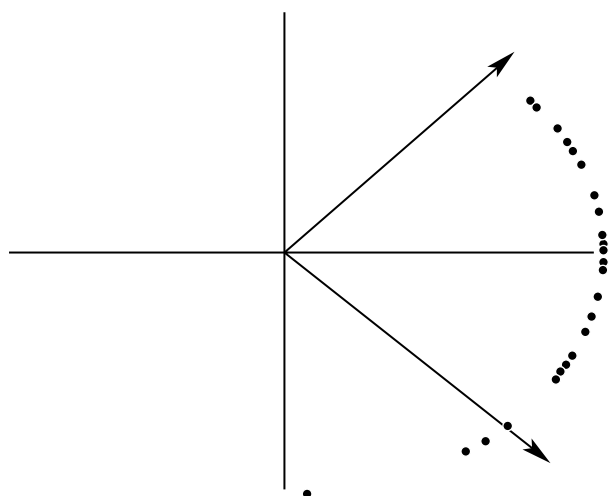
(b) Takeoff.

Figure 12. Subject plots for set A ranking for generic and takeoff conditions.

(a) Generic.

(b) Takeoff.

Figure 13. Subject plots for set B ranking for generic and takeoff conditions.

selected for use. Furthermore, although the set B dimensions were statistically significant from one another, the main effect accounted for 38.6 percent of the variance. While this percentage is good, it is not overwhelming. Finally, it perhaps should not be surprising that the flight function dimension tended to dominate the other two for the set B results, given the typical pilot training that emphasizes aviation, navigation, and communication.

The MDPREF results confirm the earlier findings that (1) pilots are in good agreement about the relative priority of the information elements they were asked to judge and (2) the high percentage of variance in the first

dimension in the earlier MDPREF analyses is reflected by the relatively tight spread of points around it in this analysis. These results, however, do underscore the potential usefulness of a second dimension, in that a good portion of subjects are using both dimensions.

Also worth noting is the consistent ratings of the four items common to both data sets. Those items were "Pilot-Initiated Request or Message," "Predicted/ Estimated Wind for Descent," "Selected Altitude," and "Speed Restriction Data." As can be seen in figures 2 and 3, these items appear in the solution space at approximately the same locations. This near identical

location would indicate that the pilots used the same attributes when measuring information similarity and that they ranked the common attributes consistently with these attributes. Likewise, these terms were similarly ranked in the prioritization task. This similarity is apparent in their common placement and shift in the spatial plots from the MDPREF analyses. These findings again underscore the utility of employing two separate data sets for validation purposes.

Cognitive Processes Across Different Contexts

One hypothesis of interest in the study reported herein was the potential for invariance of cognitive processing dimensions in different conditions. Although takeoff was the specific condition selected here and clearly differs from other conditions (such as climb, descent, and approach) in terms of the information and the actions required, it is representative of other flight conditions in the sense that specific actions need to be performed at certain times and in a certain order. The same would hold true for any other selected flight condition.

By invariance it is meant that the underlying cognitive process that is used by the pilot in accomplishing the task does not change as a function of the condition in which it occurs. This is not to say that in the case of prioritization the relative order of information cannot change as a function of context, but rather that the underlying process generating that order does not change contextually.

If, as discussed earlier, the dimensions obtained from the MDPREF analyses are representative of the pilot's cognitive processing, then the same dimensions should correlate closely across different conditions. Alternatively, dimensions reflecting different cognitive processes should still be expected to show less correlations across different conditions. Different cognitive processes should also show a dissociation within a given contextual condition. However, because the MDPREF analysis calculates principal components, and these are orthogonal, the correlation between dimensions is, by definition, zero.

This pattern of correlation was, in fact, observed in the study reported herein. The average correlation for the same dimensions between the generic and the takeoff conditions was 0.663. The average correlation for different dimensions between the generic and the takeoff conditions was −0.209. These results indicate that regardless of the particular context in which prioritization occurred, the same cognitive processes (represented here by the two dimensions) operated in a similar fashion. The fact that the different dimensions show less correlation across

contexts lends additional credence to the conditions being separate processes.

It is also important to note that, of the four stimulus elements common to both sets A and B, the relative location of three of them ("Pilot-Initiated Request or Message," "Selected Altitude," and "Speed Restriction Data") showed only minor changes across the generic and the takeoff conditions, which can be seen in figures 6 and 7. The item which showed the largest shift was "Predicted/Estimated Wind for Descent." Here the information moved toward the far left (lower perceived importance) as would be expected because it has little, if any, relevance for the takeoff condition.

Conclusions

The experiment reported herein used psychometric scaling and cluster analyses to determine how pilots mentally categorize and prioritize flight-deck information. The techniques used herein are particularly robust in that they give subjects a great degree of freedom in responding. This freedom has a clear advantage over traditional experimental designs in that subjects are allowed to provide data as they perceive the situation, without the usual constraints placed upon them in a typical experimental task (e.g., only correct or incorrect responses). It was hoped that this would afford the maximum opportunity for exploring how pilots process flight-deck information with a minimum of artificial restrictions.

A disadvantage to the methods used herein is that by providing the subjects with this freedom, individual differences among subjects may completely obscure the overall results, which was an initial concern. As pointed out in the "Results" section, it did not prove to be a problem. The results obtained for both the scaling and the clustering analyses were remarkably similar, and while there were some minor individual differences among subjects, the overall response patterns were quite similar.

Cognitive Processing of Flight-Deck Information

Results from the individual differences scaling analysis (INDSCAL) revealed three dimensions along which pilots categorized and prioritized flight-deck information. These included (1) the flight function that the information supports, (2) the perceived strategic and tactical nature of the information (referred to herein as flight action), and (3) how frequently the pilot refers to the data (referred to herein as the sample rate). These same three dimensions were observed with different subjects and different stimulus sets, which lends support for their statistical stability.

These results provide important insight into those cognitive factors involved in a pilot's processing of

flight-deck data and suggest additional information that may be collected when evaluating a new flight deck or interface. Such information could be gathered by standardizing the obtained dimensions and treating them as additional variables for early design analyses. As described in appendix A, a variety of options for using these dimensions exist.

Classification of Flight-Deck Information

Cluster results support the existence of at least four distinct categories of information. These categories include (1) aviation, (2) navigation, (3) communication, and (4) systems administration. These findings were observed across two essentially different stimulus sets and with two groups of subjects.

Based on the research discussed in the "Background" section of the paper that demonstrates that information consistent with an individual's categorization scheme is responded to quicker and more accurately, the results presented herein could have direct implications for methods of presenting flight-deck information. Most obvious is the possibility for separately displaying information within these observed categories. Such an approach may lead to performance improvements by making the information cognitively consistent by presenting it according to the mental models the crew has for the information flow. While this presentation is done for most displays on modern generation flight decks, these results could have significant bearing on how information is presented on new displays where a variety of information could be presented (such as an electronic library display). These issues will require further empirical analysis to determine precisely how such presentational issues may affect pilot performance.

Relative Priority of Flight-Deck Information Categories

The results of the MDPREF (multidimensional preference scaling) analyses provide insight into how pilots prioritize categories of information. Although, as noted in the "Background" section, most prioritization is currently done with procedural and emergency checklists, the findings presented herein may help in understanding what categories of information pilots consider to be important under nominal conditions. The current results show that the relative priorities pilots assign to various categories of information (for either generic or takeoff conditions) are

1. Aviation

2. Navigation

3. Communication

4. Systems administration

While this can provide a starting point for determining how pilots view the relative priorities of different information at a global level, it should not be taken as an absolute ordering of flight-deck information categories because of the obvious effect context can have on the relative priority of information.

Pilot Cognition Across Contexts

A central finding of the study presented herein, however, is that while the relative priorities of different information may change, depending upon the situation, the cognitive processes determining how the information is ordered are reasonably invariant with respect to context. This invariance was indicated by high positive correlations between the same MDPREF dimensions (which presumably reflect the same or similar underlying cognitive processes) and low correlations between the different MDPREF dimensions (which presumably reflect different underlying cognitive processes) across the two contextual situations.

While the dimensions obtained from the MDPREF results address relative prioritization, particularly as it relates to categories of information, further work needs to be done to determine and quantify the particular components that these prioritization dimensions are comprised.

Appendix A

Example Application of INDSCAL Dimensions

The INDSCAL dimensions obtained in the experiment discussed herein may also be used in a predictive fashion for future applications. There are two possible strategies for developing scales for research, each of which will be described below. Prior to this, however, it is useful to present the logical background for using the INDSCAL dimensions for cognitive analyses. For the following example, refer to figure A1.

Consider a pilot's cognitive state at time $t1$ to be represented at some point P1 in the three-dimensional INDSCAL plot as shown in figure A1. For example, a pilot may be engaged in a tactical navigation activity that requires relatively frequent use of data. The pilot's cognitive state at this particular point in time has been represented in figure A1 as P1.

Assume now that a tactical aviation activity arises (such as a request from ATC to climb to a flight level of 35 000 ft). The question becomes, can the pilot engaged in the long-term navigation task effectively deal with this new task, or will imposition of the new task create a dis-

ruption in performance? This question bears on a central issue in multiple-task performance research—namely, the extent to which different tasks may interfere, or conversely, facilitate each other.

In the case described above, the issue of what resources tasks require for adequate performance emerges. First, the addition of a task from the same high-level category (e.g., an aviation task with an aviation task) may lead to superior performance, if it is assumed that the subject is primed for performing such a task because they are currently performing tasks from the same area. Alternatively, the addition of a task from within the same category might lead to an overload situation where performance actually declines because both tasks could be assumed to require the same resources. Both empirical questions require further analysis.

An analogous series of questions emerges when the addition of tasks from outside the same category is considered. One question is, can a subject adequately time-share two tasks from different functions (e.g., an aviation task with a navigation task) if they may be presumed to require different resources? Another question is, might the tasks interfere with one another?
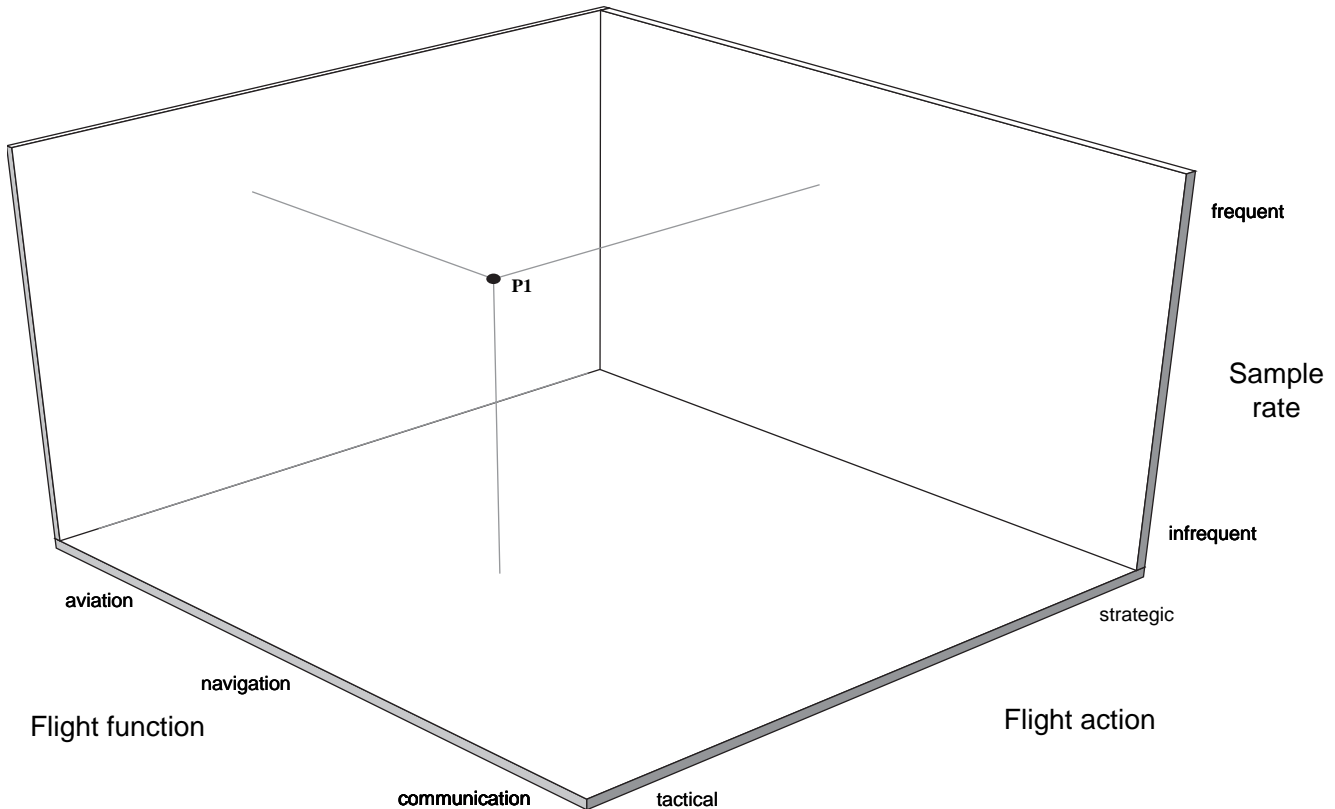

Figure A1. Hypothetical example of pilot's cognitive state using INDSCAL dimensions.

This example illustrates how the INDSCAL dimensions may be used for evaluating flight-deck designs. Because the pilot may be represented at any point in the flight by three INDSCAL dimensions and these dimensions relate to the pilot's processing of flight-deck information, the issue of using the dimensions to determine whether subsequent tasks will be compatible with ongoing activities becomes quantifiable. Once appropriate validation has taken place, such scales could become extremely useful for flight-deck design. At least two possibilities for validation present themselves.

First, the dimensional scales could be part of a traditional task analysis, wherein a record of the information the operator used to accomplish the task could be accumulated on each of the three dimensions. Typically, task analyses examine the amount of time required to perform a task and the total amount of time available for the given period. While overload (the individual has too little time to perform too many actions) can be detected in such a situation, and represents a proxy for workload, little information is conveyed about the pilot's cognitive processing from this methodology. Dimensional information would provide, at a particular point in the timeline, a rating for the pilot. The only requirement would be that the analyst have knowledge of the flight-deck information used by the pilot for a given task. Currently, this is not a problem because as part of any detailed task, the actions (including display monitoring and MCDU interactions) are explicitly spelled out through scenario development with experienced pilots. If scale measures could be calibrated to identify high and low cognitive workload levels, such scale values could be used to estimate potential cognitive over or underload situations.

In a second method, these scales could be used jointly. This approach would be the more difficult of the two. Here, instead of presenting a task by task estimate along each scale, the task loads could be legitimately combined mathematically and theoretically. This combining would probably be best accomplished by using conjoint measurement techniques (Louviere 1988). With this approach, the analyst would be able to examine the dimensional scores for (1) a particular task or series thereof, (2) a defined mission segment (such as takeoff), and/or (3) the entire mission.

The above conditions suggest that the most logical approach for using the derived INDSCAL dimensions is as scales for pilots to evaluate and rate flight-deck information. Three INDSCAL dimensions (flight function, flight action, and sample rate) and their operational definitions could be provided to pilots. Once becoming familiar with these scales, pilots could then rank a complete list of flight-deck information (such as the list used to construct the data sets in the study presented herein) along those three dimensions.

Appendix B

Categorization and Prioritization Questionnaire

I. During the course of a flight you receive a variety of different types of information. Can you provide a list of general categories into which you place these sources of flight deck information?

II. What factors would you consider to be of the most importance for prioritizing flight deck information and of the least importance?

III. Are there any types of flight deck information which you would consider to have absolute levels of priority (that is, you would attend to these things over any other information or, conversely, ignore)?

IV. In normal flight, how do you go about prioritizing information?

References

Abbott, Terence S. 1993: *Functional Categories for Future Flight Deck Designs*. NASA TM-109005.

Anon. 1988: S*AS/STAT® User's Guide, Release 6.03 Edition,* SAS Institute Inc.

Anon. 1992: Boeing Model 747-400—Operations Manual, Volume 2. No. D6-30151-401, Boeing Commercial Airplanes.

Ashby, F. Gregory, ed. 1992: *Multidimensional Models of Perception and Cognition*. Lawrence Erlbaum.

Ashby, F. Gregory; and Maddox, W. Todd 1993: Relations Between Prototype, Exemplar, and Decision Bound Models of Categorization. *J. Math. Psychol.,* vol. 37, pp. 372–400.

Braune, Rolf J.; Hofer, Elfie F.; and Dresel, K. Michael 1991: Flight Deck Information Management—A Challenge to Commercial Transport Aviation. *Proceedings of the Sixth International Symposium on Aviation Psychology,* Volume 1, R. S. Jensen, ed., Ohio State Univ., pp. 78–84.

Carroll, J. Douglas 1972: Individual Differences and Multidimensional Scaling. *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences*, Roger N. Shepard, A. Kimball Romney, and Sara Beth Nerlove, eds., Seminar Press, pp. 105–155.

Carroll, J. Douglas; and Chang, Jih-Jie 1970: Analysis of Individual Differences in Multidimensional Scaling Via an N-Way Generalization of Eckart-Young Decomposition. *Psychom.,* vol. 35, no. 3, pp. 283–319.

Cliff, Norman 1987: *Analyzing Multivariate Data*. Harcourt Brace Javanovich, Inc.

Dornheim, Michael A. 1992: Changes Hike MD-11 Training Success. *Aviation Week & Space Technol.,* vol. 137, no. 21, pp. 93–95.

Edwards, Ward 1987: Decision Making. *Handbook of Human Factors,* Gavriel Salvendy, ed., John Wiley & Sons, pp. 1061–1104.

Glass, Arnold Lewis; and Holyoak, Keith James 1986: *Cognition.* Second ed., Newberry Award Records, Inc.

Kruskal, Joseph B; and Wish, Myron 1978: *Multidimensional Scaling*. Sage Publ.

Louviere, Jordan J. 1988: *Analyzing Decision Making—Metric Conjoint Analysis*. Sage Publ.

Nosofsky, R. M. 1984: Choice, Similarity, and the Context Theory of Classification. *J. Exp. Psychol.: Learn, Mem., & Cognit.,* vol. 10, pp. 104–114.

Nosofsky, R. M. 1986: Attention, Similarity, and the Identification-Categorization Relationship. *J. Exp. Psychol.: Gen.*, vol. 115, pp. 39–57.

Nosofsky, Robert M. 1992: Similarity Scaling and Cognitive Process Models. *Annu. Rev. Psychol.*, vol. 43, pp. 25–53.

Posner, M. I.; and Keele, S. W. 1970: Retention of Abstract Ideas. *J. Exp. Psychol.*, vol. 83, no. 2, pp. 304–308.

Reed, Stephen K. 1972: Pattern Recognition and Categorization. *Cogn. Psychol.*, vol. 3, pp. 382–407.

Ricks, W. R. 1994: *PWC—Pairwise Comparison Software: Software Program for Pairwise Comparison Task for Psychometric Scaling and Cognitive Research*. LAR-15143.

Ricks, Wendell R.; Jonsson, Jon E.; and Rogers, William H. 1994: Cognitive Representations of Flight-Deck Information Attributes. *Int. J. Aviat. Psychol.*, vol. 4, no. 1, pp. 65–83.

Ricks, Wendell R.; Abbott, Kathy H.; Boucek, George; Jonsson, Jon; and Rogers, William H. 1991: Information Management for Commercial Aviation—A Research Perspective. *Proceedings of the Sixth International Symposium on Aviation Psychology*, R. S. Jensen, ed., Ohio State Univ., pp. 60–65.

Rips, L. J.; Shoben, E. J.; and Smith, E. E. 1973: Semantic Distance and Verification of Semantic Relations. *J. Verbal Learn. & Verb. Behav.*, vol. 12, no. 1, pp. 1–20.

Romesburg, H. Charles 1984: *Cluster Analysis for Researchers*. Wadsworth, Inc.

Sarter, Nadine B.; and Woods, David D. 1994: Pilot Interaction With Cockpit Automation II: An Experimental Study of Pilots' Model and Awareness of the Flight Management System. *Int. J. Aviat. Psychol.*, vol. 4, no. 1, pp. 1–28.

Shepard, Roger N. 1987: Toward a Universal Law of Generalization for Psychological Science. *Science*, vol. 237, pp. 1317–1323.

Wiener, Earl L. 1989: *Human Factors of Advanced Technology (Glass Cockpit) Transport Aircraft.* NASA CR-177528.

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
August 1995

3. REPORT TYPE AND DATES COVERED
Technical Paper

4. TITLE AND SUBTITLE
Cognitive Models of Pilot Categorization and Prioritization of Flight-Deck Information

5. FUNDING NUMBERS
WU 505-64-13-22

6. AUTHOR(S)
Jon E. Jonsson and Wendell R. Ricks

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Langley Research Center
Hampton, VA 23681-0001

8. PERFORMING ORGANIZATION REPORT NUMBER
L-17463

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA TP-3528

11. SUPPLEMENTARY NOTES
Jonsson: McDonnell Douglas Aerospace-West, Long Beach, CA; Ricks: Langley Research Center, Hampton, VA.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified–Unlimited
Subject Category 01
Availability: NASA CASI (301) 621-0390

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In the past decade, automated systems on modern commercial flight decks have increased dramatically. Pilots now regularly interact and share tasks with these systems. This interaction has led human factors research to direct more attention to the pilot's cognitive processing and mental model of the information flow occurring on the flight deck. The experiment reported herein investigated how pilots mentally represent and process information typically available during flight. Fifty-two commercial pilots participated in tasks that required them to provide similarity ratings for pairs of flight-deck information and to prioritize this information under two contextual conditions. Pilots processed the information along three cognitive dimensions. These dimensions included the flight function and the flight action that the information supported and how frequently pilots refer to the information. Pilots classified the information as aviation, navigation, communications, or systems administration information. Prioritization results indicated a high degree of consensus among pilots; while scaling results revealed two dimensions along which information is prioritized. Pilot cognitive workload for flight-deck tasks and the potential for using these findings to operationalize cognitive metrics are evaluated. Such measures may be useful additions for flight-deck human performance evaluation.

14. SUBJECT TERMS
Human factors; Cognition; Pilot performance; Categorization

15. NUMBER OF PAGES
32

16. PRICE CODE
A03

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified

20. LIMITATION OF ABSTRACT