Exploiting Archival Data to Identify CRM Training Needs for C-130 Aircrews

Robert T. Nullmeyer
Air Force Research Laboratory
Mesa Arizona

V. Alan Spiker
Anacapa Sciences
Santa Barbara, California

Given the time, money and effort being invested to satisfy Crew Resource Management (CRM) training requirements, there is surprisingly little empirical data to guide CRM course content. In military programs, CRM training is often organized around a set of elements that are listed in service-wide training regulations, with roughly equivalent attention being paid to each area. If some CRM processes are more closely related to mission outcomes than others, the value received from CRM training might be increased through greater focus on areas of greatest need. Military and commercial aviation research findings and recommendations led the Air Force to require a more data-driven approach to establish specific behaviorally anchored CRM training objectives. The Air Mobility Command, the C-130 Aircrew Training System Program Office, Lockheed Martin, and The Air Force Research Laboratory formed a partnership that is now conducting a series of studies and analyses to identify and prioritize CRM training requirements for C-130 aircrews. This paper discusses two analyses of existing data: (1) instructor observations recorded in student training folders during mission qualification training, and (2) mishap report narratives and associated databases of causal and contributing human factors that are maintained by the Air Force Safety Center. In both cases, the original data were generated for other purposes, but were made available for these CRM analyses. Each proved to be a fertile source of insight regarding specific CRM behaviors that need to be considered in redesigning CRM instruction. We describe these two data sets, our analytical approach, trends identified, and implications for CRM training. We anticipate that these analyses will contribute to the development of observable CRM training objectives that will, in turn, enable compelling transfer of training assessments of resulting changes in CRM training.

About the Authors

Robert T. Nullmeyer is a research psychologist with the Warfighter Training Research Division of the Air Force Research Laboratory at Mesa, AZ (AFRL/HEA), where he has been since 1981. He has conducted research on training system evaluation, and simulator training effectiveness. He earned a Ph.D. in Experimental Psychology from the University of New Mexico.

V. Alan Spiker is a principal scientist at Anacapa Sciences, Inc., Santa Barbara, CA. He is responsible for human factors research associated with advanced technology and has conducted field studies on mission planning and command and control systems for all branches of service. He has a Ph.D. in Experimental Psychology from the University of New Mexico and has been with Anacapa Sciences since 1982.

Exploiting Archival Data to Identify CRM Training Needs for C-130 Aircrews

Robert T. Nullmeyer
Air Force Research Laboratory
Mesa Arizona

V. Alan Spiker
Anacapa Sciences
Santa Barbara California

INTRODUCTION

Over the past two decades, virtually every aviation training program has adopted Cockpit/Crew Resource Management (CRM) training or its close relative, Aircrew Coordination Training (ACT). Despite this widespread support of CRM instruction by aviation training organizations, there is a surprising lack of consensus regarding some fundamental properties of this instruction including training objectives, content, and strategies for effective delivery. Salas, Prince, Bowers, Stout, Oser, and Cannon-Bowers (1999) reported considerable variability across CRM programs, with duration ranging from one hour to two weeks and program content addressing varying combinations of interaction styles, stress reduction, and automation issues in addition to the more standard workload management, advocacy, and situational awareness elements. Some programs appear to be designed to facilitate attitude change, others focus on skill development, and still others have no discernible basis.

Helmreich, Merritt, and Wilhelm (1999) documented five distinct generations of CRM in commercial aviation since its introduction at United Airlines in 1981. Programs representing various generations are in use today. Early generations were seminar-based and psychological in nature. They addressed general management concepts such as leadership and emphasized correcting deficient behaviors in individuals such as lack of assertiveness on the part of flight engineers or authoritarian behavior by airline captains. In later generations, the scope expanded beyond individuals to address team dynamics in the cockpit, sometimes including other members of the flight environment such as flight attendants and maintenance personnel. The focus shifted away from general concepts toward specific behaviors that could be used by pilots to function more effectively. The introduction of the Advanced Qualification Program (AQP) in 1993 furthered this evolution of CRM by requiring detailed analyses of training requirements for each aircraft, programs to address human factors (CRM) issues in each aspect of training, and integration of CRM concepts into technical training. AQP also required the use of full mission simulation for Line Oriented Flight Training (LOFT) and Line Operational Evaluation, adding demonstration, practice, and feedback to the CRM training tool kit.

CRM instruction in the Air Force followed a similar evolution. Air Force Instruction (AFI) 36-2243 established the requirement for all Air Force aviators to receive general CRM familiarization training. A variety of factors contributed to the development of a few core training packages that were then exported to multiple operational communities--a practice that was also somewhat common in commercial aviation at the time. Helmreich, et al. reported that these imported courses had consistently less impact in the new airline than the original had when it was delivered to the community for whom it was originally developed. Similarly, Silverman, Tourville, Spiker, and Nullmeyer (1997) documented several problems with the delivery of generic CRM courseware, including little overlap between CRM course content and specific CRM behaviors that were demonstrated to be highly related to mission outcome.

Based on these research findings and strong operational command input, AFI 11-290 was rewritten and then implemented in 1998 to update Air Force CRM training policy. This AFI defines CRM as "the effective use of all available resources--people, weapon systems, facilities, equipment and environment--by individuals or crews to safely and efficiently accomplish a mission or task." Six core curriculum areas are specified--situational awareness, crew coordination/flight integrity, communication, risk management/decision making, task management, and mission planning/debrief. Mission-specific CRM training programs are required along with training objectives that are tailored to the knowledge and skill level of the aircrew member.

In the following sections, we describe two analyses that were done to help identify and prioritize the CRM behaviors to be addressed in training based on a criterion of probable positive impact on subsequent aviator performance. The analysis of instructor comments in MC-130P student records was recently completed to support a substantial redesign of CRM training at the 58th Special Operations Wing, Kirtland AFB NM. The analysis of Air Force Safety Center mishap data is part of a new effort to modernize CRM instruction at the 314th Tactical Airlift Wing, AR. The Air Education and Training Command sponsored the MC-130P student records analyses. The Safety Center data analyses are supported by the C-130 Aircrew Training System Program Office. Both projects were part of larger research projects that included substantial documentation and analyses of CRM behaviors and mission performance data during Mission Oriented Simulator Training (MOST), allowing triangulation across multiple data sources.

MC-130P TRAINING RECORDS--SEARCHING FOR A SENSITIVE INDEX OF CRM PROFICIENCY

Student Grades

This analysis was derived from training records obtained from the 58th Special Operations Wing (SOW), Kirtland AFB, NM. For each academic, simulator, and flightline training session, the instructor assigns a letter grade (P=proficiency advance, E=exceptional, S=satisfactory, T= more training required, U=unsatisfactory, I=incomplete due to weather or maintenance) documented in the student's Form 15, "Aircrew Training Record." To be a useful source of proficiency data, grades must vary across students, else one cannot infer variations due to program variables or training interventions. However, in our preliminary analysis of MC-130P student records, fewer than 2% of student grades in the Form 15 were other than S.

MQ training progresses through a conversion simulation mission, two day-tactical simulations, and four night tactical simulation missions. This is followed by flightline training, where students fly two or three day tactical missions and then seven or eight night tactical missions. The conditions become more difficult as training progresses, with darker night skies and more demanding mission profiles. To keep track of this progression, instructors fill out a Form 14, "Aircrew Training Progress Record," after each simulation and flightline mission. This form provides a set of pre-printed required proficiency levels (RPLs) for the training "events" associated with that mission profile. Events are task-based, and include such activities as airdrop checklist, simulated engine failure, NVG operations, and the like. Performance is graded on a four point scale (1=extremely limited, 2=partially proficient, 3=competent, 4=highly proficient) as is knowledge (ABCD). As the student progresses through training, the preprinted RPLs for the training events in each mission profile become more stringent (e.g., C3, D3 rather than A1 or B2).

The instructor crosses off RPLs on the pre-printed form for each training event that is accomplished. If a student exceeds or fails to meet an RPL, the instructor must write in the actual level of performance or knowledge demonstrated. But in practice, instructors rarely note deviations from the RPL because they do not wish to have a student's deficiencies noted in his/her permanent record. When substandard performance arises, the instructor remediates the student "off-line" in a way that is not reflected in the student's recorded grade (Bruce, Killion, Rockway, & Povenmire, 1991). As a result, analysis such as aggregating the number of events failing to meet the RPL for that mission profile will not yield sensitive measures of proficiency since most entries are the unaltered RPLs.

Instructor Comments

Instructors provide written comments concerning student performance for each simulator and flight session. These are documented in the training folders on Form 13, "Training Comments Record." The comments are unstructured and not necessarily tied to the required items in Form 14. While unstructured, these comments provide a wealth of information that is a potentially rich source of proficiency information. In the comments, instructors are free to note their reservations regarding a student, knowing that their remarks are not reflected in the student's recorded grade. Moreover, the instructor can go over the comments with the student after a training mission, using it as a teaching or debriefing aid.

In preliminary analyses of training records for each MC-130P crew position (Spiker, Tourville, & Nullmeyer, 1999), we determined that a substantial number of instructor comments are recorded and, when aggregated across missions, can be reliably classified into positive and negative cases. Moreover, the comments can be sorted into functional categories (e.g., crew coordination, equipment knowledge) characteristic of each crew position. Comments are quite specific (e.g., "missed several radio calls," "must keep checklist flowing to ensure proper crew responses," “need more positive continuous guidance to pilot”) and appeared to yield valuable insights into areas where student proficiency is strong or weak. To the extent that these comments are recorded routinely and comprehensively, they can be content-analyzed, aggregated, and quantified to yield data-based assessments of student proficiency.

Encouraged by the richness of instructor comments, we had two subject matter experts (SMEs) perform comprehensive, independent reviews of MC-130P Mission Qualification (MQ) training records. Both SMEs were experienced trainers, one in airborne command & control, and the other in SOF fixed wing aircraft. The purpose of this initial review was to determine if SMEs could use the instructor comments in the MQ training folders to derive consistent and reliable estimates of student proficiency.

The records were taken from aircrews receiving MQ training during the first part of 1998. Representative records were sampled from the five MC-130P crew positions trained in MQ—pilot, navigator, flight engineer (FE), communication systems operator (CSO), and loadmaster (LM). Four records were reviewed for each crew position, for a total of 20 records. Once demographic data were obtained from each record, the student’s name was purged from our research files and his/her data were not used for any purposes other than research.

To structure their review, the two SMEs highlighted all instructor comments relevant to CRM topics. These comments were then paraphrased and transcribed onto a four-page, structured Training Record Evaluation Worksheet. A portion of this worksheet is depicted in Figure 1, along with

MISSION PLANNING / DEBRIEF—*Includes pre-mission analysis and planning, briefings, ongoing mission evaluation, and post-mission debrief. Considers general knowledge and use of specific mission planning tools and/or operations techniques.*									
	Basis for Student Proficiency Assessment	**Mission**	**Rating**						
1.	Good understanding of SOFPARS mission planning (pilot)	DT-1	1	2	3	X	5	ID
2.	Excellent permission planning effort + ground ops (nav)	NT-1	1	2	3	4	X	ID
TASK MANAGEMENT—*Includes establishing priorities, managing multiple tasks, adaptability, flexibility in responding to task-overload/underload, complacency, management of automation, use of resources, checklist discipline, and SOPs.*									
1.	Responded to checklists in appropriate tone (pilot)	DT-1	1	2	X	4	5	ID
2.	Cross check was slow resulting in fast Rz (nav)	NT-2	1	X	3	4	5	ID
3.	Skips over steps in checklists (nav)	NT-5	X	2	3	4	5	ID
SITUATIONAL AWARENESS—*Includes knowledge and skills for preventing the loss of situational awareness, recognizing the loss of situational awareness, and techniques for recovering from the loss of situational awareness.*									
1.	Unsure of what calls to make at what time (CSO)	NT-1	1	X	3	4	5	ID
2.	Failed to accelerate to 150 kts while Nav directed the helos (pilot)	NT-2	1	X	3	4	5	ID
CREW COORDINATION / FLIGHT INTEGRITY—*Includes command authority, leadership, assertiveness, conflict resolution, hazardous attitudes, behavioral styles, legitimate avenues of dissent, and team building.*									
1.	Excellent support for Pilot and Nav (CSO)	SDT-2	1	2	3	X	5	ID
2.	Slow to copy and relay info to crew (CSO)	SNT-2	1	X	3	4	5	ID
COMMUNICATION—*Knowledge of common errors, cultural influences, and barriers (rank, age, exper., position), attitude/ motivation. Skills include listening, feedback, precision and efficiency of communication with all crewmembers and agencies.*									
1.	Need to be more assertive (FE)	NT-2	1	X	3	4	5	ID
2.	Not afraid to question pilot or nav (LM)	DT-X	1	X	3	4	5	ID
3.	Missed several radio calls (CSO)	NT-5	1	X	3	4	5	ID
RISK MANAGEMENT / DECISION-MAKING—*Includes risk assessment, the risk management process, tools, breakdowns in judgment and flight discipline, problem-solving, evaluation of hazards, and control measures.*									
1.	Good air sense (pilot)	SDT-3	1	2	3	X	5	ID
2.	Slow to get favorable bank angle during simulated engine loss (pilot)	NT-2	1	X	3	4	5	ID
Tactics Employment—*Includes knowledge and skills necessary to avoid or minimize threat detection or exposure, and to successfully coordinate complex mission events and multiple mission objectives. Also includes tactical maneuvers, threat-related mission planning, and inflight changes/re-planning in response to evolving tactical circumstances.*									
1.	Excellent evasive maneuver for threat calls (nav)	SNT-2	1	2	3	4	X	ID
2.	Have to use expendables to manage them (nav)	SNT-4	X	2	3	4	5	ID

Figure 1. Example instructor comment data and SME ratings from Training Record Evaluation Worksheet

example data. The Worksheet is organized around the six CRM areas covered in Air Force Instruction 11-290. To these six we added a seventh area, Tactics Employment (TE), defined as the "knowledge and skills necessary to avoid or minimize threat detection or exposure, and to successfully coordinate complex mission events and multiple mission objectives." Each paraphrased comment was placed in the relevant CRM category, and then rated on a five-point scale (1=poor, significantly below expectations; 2=marginal, less than desired; 3=standard; 4=very good, above expectations; 5=exceptional, a model of high level skill; ID=insufficient data).

We found in our preliminary analysis that instructor comments tended to follow a number of distinct topics or "themes," and these themes could be linked to individual CRM areas. Table 1 lists the most prominent themes for the six CRM areas and TE. SMEs used these listings as an intermediate aid in assigning instructor comments to the CRM areas shown in Figure 1. Once the comments were transcribed and rated on a five-point scale, a summary rating for that CRM category was assigned. Finally, an overall proficiency rating was assigned for each student.

Results

Inter-Rater Reliability. A critical goal of the analysis was to determine if the two SMEs were consistent in assigning student proficiency ratings. We computed the correlation between the two sets of 20 overall proficiency ratings (one per student) from the SMEs. The correlation between these ratings was .81, which exceeds the .80 value typically noted (Cronbach, 1990) for acceptable inter-rater reliability. Looking at the ratings themselves, the two raters produced identical overall ratings for 16 of the students; their ratings differed by only 1 scale value for the other four students. We conclude that SMEs can reliably assign proficiency ratings to the comments that instructors place in student grade folders.

Quantitative. Having established the inter-rater reliability associated with this rating process, we had one SME extend his review to include all MC-130P student records from 1998, for a total of 87 students. We then examined the absolute values of the overall ratings, to determine if there is sufficient variation across students to support inferences concerning CRM areas in need of improvement. Of the 87 records reviewed by our SME, more than one-third received a rating other than "3." In contrast, only 3.1% and 14.6% of simulator and flights, respectively, resulted in a grade other than "S." For the flight evaluation, all 87 students received a Q-1 qualification level and only 1.4% of flights produced an evaluation other than a "B."

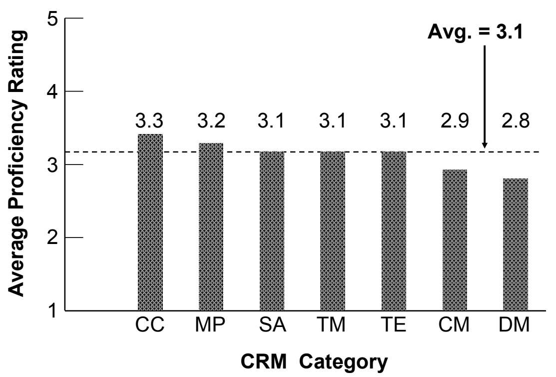
The sensitivity in the comment data is apparent when the overall proficiency ratings are broken down by CRM category, as shown in Figure 2. Computing the average rating variability within each category, we determined that a scale difference of .14 can be considered statistically meaningful (Hays, 1973). On that basis, we see that Mission Preparation and Crew Coordination

Table 1. Behavior Themes Associated with Each CRM Category

CRM Category	Behavior Theme
Mission Planning/Mission Evaluation	Briefings, route study, mission analysis, mission preparation, knowledge of MP tools and techniques, on-going mission analysis during execution TOLD data
Situational Awareness	Threat awareness and avoidance, recognize unusual situations ("good catch"), staying ahead or falling behind aircraft, getting lost/confused, calculating position incorrectly
Crew Coordination/Flight Integrity	Leadership, assertiveness, formation coordination, coordination with other crewmembers, backing up crew, staying ahead of crew, need to be directive, speak up, taking charge, share info with rest of crew
Communication	Precise and efficient comm., terminology, missing radio calls, stepping on calls, comm procedures, talking too much or too little, good alt/speed calls, terrain descriptions
Decision Making/Risk Management	Risk assessment, evaluation of hazards, safety issues, breakdown of judgment or discipline, knowledge of emergency equipment, response to and avoidance of threats, managing mission profile, dealing with fluid sits, keeping up with mission changes, good decisions, poor aircraft control, adapt to changing conditions
Task Management	Task prioritization, time management, SOPs, slow at performing duties, response to task overload/underload, equipment/procedures knowledge, monitor instruments/information, cross check instruments, checklist discipline, aircraft handling
Tactics Employment	Threat briefs, aerial refueling techniques, terrain following maneuvers, airdrop procedures, expendables, radar interpretation, and knowledge of secure communications

received significantly above-average ratings (3.24 and 3.28 respectively), with Decision-Making and Communications (2.83 and 2.89 respectively) significantly below average. The overall rating was 3.10.

Figure 2. Average Student Proficiency Rating by CRM Category



We also examined the overall proficiency ratings for each crew position. Using the .14 scale difference as above, we found that the FE received significantly higher ratings (mean=3.4) compared to the pilot (3.2), navigator (3.1), and CSO (3.1). The loadmaster received a significantly lower average rating (2.9). Follow-up interviews with 58 SOW instructors confirmed this breakout, as they cited the high quality of FE training provided there as well as the problems posed by the lack of a simulator seat for the LM.

Qualitative. Having identified the CRM areas that stand out statistically, we then examined the associated instructor comments to pinpoint specific areas where present CRM behaviors are strong and weak. In performing this analysis, it should be noted that most comments have two aspects: evaluative and directive. The evaluative component gauges student proficiency in the commented area, and is typically represented by an adjective, such as "good" mission planning, "excellent" mission briefing, or "weak" situation awareness. The directive aspects of each comment let us extract the specific crew behaviors that were either deficient or laudable. By directive, we mean such comments as "slow to prepare brief," "needs to think further ahead of the aircraft," or "missed radio calls." These comments are usually given to the student as verbal feedback during the training session, to promote immediate improvement or to reinforce some essential skill. Over the long term, the content of these comments can be collected, analyzed, and folded back into an improved training curriculum as a set of target behaviors.

To illustrate, we reviewed the compiled set of instructor comments in Mission Preparation and Crew Coordination to ascertain why these categories stood out as above average. For the most part, instructors were complimenting students on such aspects as "thorough" planning, "concise" briefings, and "good backing up crewmembers." These areas are emphasized in the present training, and seem to have been internalized by the students. Yet there were also negative comments indicating areas in need of improvement, such as navigation and leadership. Examples include the need for: greater annotation of significant terrain features on maps, discussing more obstacles in the low-level brief, and taking firmer control of the crew.

Turning to the weaker CRM areas, Decision Making exhibited a wide assortment of deficiencies that primarily involved pilots and navigators. A major deficiency entailed slow reactions to conditions requiring more rapid judgment, such as initiating emergency procedures, responding to loss of engine, turning to final approach, joining-up during aerial refueling, correcting the flight profile, and breaking off formation during the onset of instrument meteorological conditions. A host of Communication problem areas was also exposed for all crewmembers. These include missing air traffic control calls (pilot), weak procedural terminology (LM), as well as the need to: provide more guidance to pilot (navigator), break in when necessary (CSO), and be more assertive (FE).

MISHAP REPORT ANALYSES

The safety organizations in each of the services expend considerable effort investigating the human factors aspects of mishaps--with good reason. Luna (2001) reported that over 60% of Air Force Class A mishaps (Over $1 million damage and/or fatality) involved human factors as a major contributor or causal factor from 1991 to 2000. Human factors were involved in over 90% of fatal aviation mishaps in that time period. The services maintain detailed databases of human factors related to these mishaps, and grants access to agencies involved in aviation safety, mishap, and injury prevention.

Overall Findings

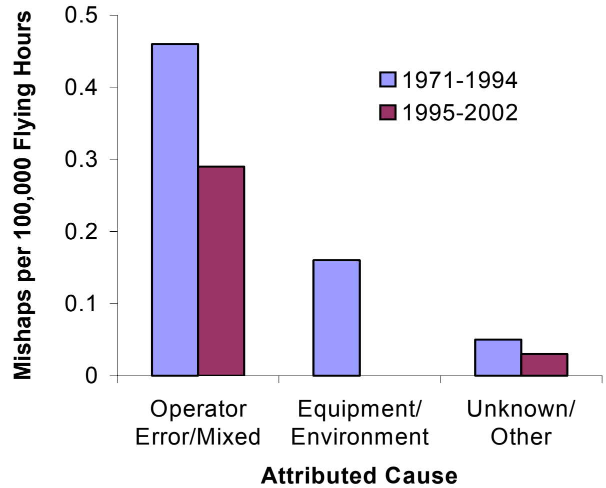
The Life Sciences Branch, Aviation Safety Division of the Air Force Safety Center provided access to their databases and worked with us to identify CRM elements that caused or contributed to Class A mishaps in the Air Force C-130 community. A Safety Center statistical summary table listed 62 Class A mishaps from 1971 through spring 2002. A *Microsoft Word* file was provided by the Safety Center that contained summaries of each of these mishaps including narratives, causal and contributing factors, recommendations, and information such as date, day/night, and phase of operation. From this file, we found that 34 of the 62 Class A mishaps were attributed solely to human factors. In eleven additional mishaps, human factors were causal along with other factors, usually equipment failure. We combined these two data sets, yielding a total of 45 human factors-related mishaps. Of the remaining mishaps, twelve were attributed to equipment or environmental factors only, and five cited unknown or other causes.

Rates per 100,000 flying hours for all C-130 Class A mishaps are depicted in Figure 3 for the periods 1971 through 1994 and 1995 through 2002. These time periods were chosen because CRM training was introduced in early 1995. Several trends can be seen. First, as has been shown in multiple aviation settings, operator error is a causal factor in most of these mishaps (73%). Second, both operator error and equipment failure rates were lower in the time period from 1995 to the present time. These trends are consistent with broader Air Force mishap trends (Luna, 2001).

Mishaps in commercial aviation occur most frequently during takeoff/climb or approach/landing phases of flight. Of some note, C-130 mishaps did not follow this pattern. A larger proportion of C-130 mishaps occurred during cruise portions of the mission (44%) than during either taxi out through takeoff climb (20%) or final approach, landing, and taxi after landing (36%).

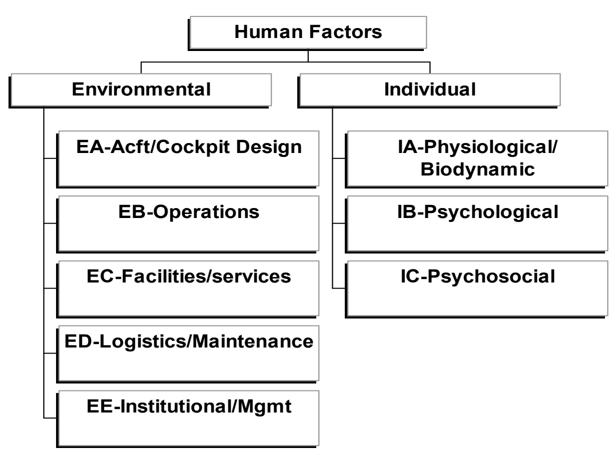
Our next step was to consider CRM-related human factors in greater detail. The Safety Center provided two human factors databases in Excel format. One covered the time period from 1971 through 1995, and a second, 1996 through 2000. For our purposes, the two databases yielded comparable data. To create the human factors lists in these databases, an analyst reviews relevant material, including interviews and reviews of medical, training, and administrative records. A human factors taxonomy provides the structure for documenting causal or contributing human

Figure 3. C-130 Class A Mishaps



factors. The analysis logic tree starts with two major branches--environmental or individual factors. These two branches are further divided into the areas shown in Figure 4. Each of these areas is further divided into sub-areas. For example, the area labeled "Operations" is coded EB for area B under Environmental Factors, which is further divided into preparation, cockpit/crew resource management, procedural guidance/publications, and mission demands. Finally, several elements comprise each sub-area. In this taxonomy, cockpit/crew resource management is coded EB2--sub-area number two of the four categories under the EB (Operations) area. Finally, several elements comprise each sub-area. EB204 (Leadership), for example, is element 04 of the EB2 subarea, one of six elements. Of 383 total elements, roughly 50 correspond to one of the CRM areas.

Figure 4. Human Factors Taxonomy



AFI 11-290 definitions are shown in italics in Figure 1 for the six CRM elements specified for Air Force CRM training. Based on these definitions, the taxonomy was searched for exemplars of each CRM category. Multiple elements were identified for each as listed in Table 2. We then searched the human factors databases, recording presence or absence of each CRM-related element within each of the 45 human factors-related mishaps.

Table 2 shows the number of mishaps in which a specific element was included in the human factors databases. The data set was split into two time periods to enable us to compare rates before and after CRM instruction was introduced at the C-130 formal school in early 1995. Thirty-nine mishaps comprise the data set from 1971 through 1994 and six additional mishaps comprise the 1995-2002 data set. Consistent with the lower overall mishap rates since 1990, rates improved for most individual CRM categories, as shown in Figure 5. The exception was communication, where the rate increased slightly. A Chi-Square analysis revealed that factors cited were not evenly distributed across the six CRM categories (Chi-Square=73.72, df=5, p<.001). Crew coordination, situational awareness, and decision making/risk analysis elements were cited more frequently. Communication elements were the least frequently cited. The National Transportation Safety Board also reported a lack of evenly distributed mishap causes in commercial aviation, with communication also being relatively infrequently cited.

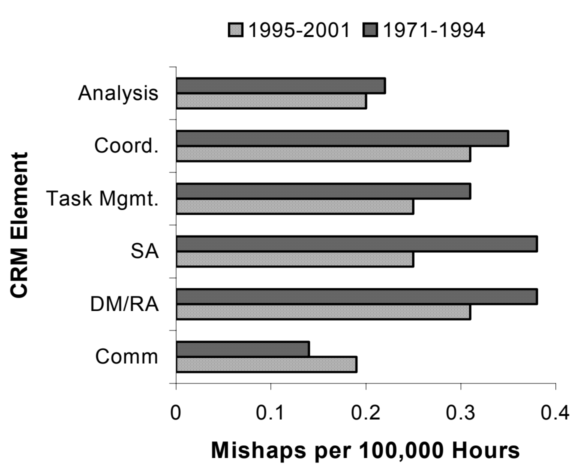


Figure 5. Mishap Rates for CRM Factors

Mishap Trends within CRM Categories

Keeping in mind that there have been only six Class A mishaps since the introduction of CRM training in 1995, there appear to be some consistent behaviors shared across those mishaps. In the following paragraphs, we summarize those behaviors in each CRM area and compare them against historic trends.

Among the ***mission planning*** elements, flight briefing dominated the list of causal and contributing factors in 1971 through 1995, being cited in almost half of the Class A mishaps. This rate continued in 1995 through 2002 with three of the six mishaps citing shortfalls in the flight briefing. However, flight planning and inflight analysis shortfalls appear to be increasing. In four of the six more recent mishaps, findings included insufficient mission planning on the part of the mishap crew. Inflight analysis is cited in 3 of the 6 mishaps. Recommendations included developing a pre-mission process that provides all appropriate mission planning information, acknowledging that current planning procedures are written primarily for pre-mission planning. The need was established to enhance procedural guidance for inflight planning, and establishing procedures and training for in-flight use of the Portable Flight Planning System (PFPS).

Crew coordination continues to be a frequently cited CRM area for causal factor in Class A mishaps, being cited in five of the six most recent mishaps. Excessive motivation and overconfidence were frequently mentioned historically. In the past few years, mishap rates in these areas have decreased, but subordinate style/copilot syndrome and complacency rates have increased. The former was cited in two-thirds of the Class A mishaps since 1995, and the latter in half of the Class A mishaps in the same time period. Subordinate style is a CRM factor in the taxonomy that applies when an individual has the basic belief that other crewmembers have the situation under control and are looking out for their best interest. Complacency refers to reduced consciousness due to an attitude of overconfidence or under-motivation.

Within the element of ***task and time management***, necessary action-delayed, has improved relative to historic rates, dropping from 55% to 33%. Cognitive task oversaturation (the quantity of information to process exceeds a person's cognitive or mental resources), however, played a role in all but one of the most recent six Class A mishaps.

Table 2. Frequency of CRM Elements Cited as Causal or Contributing Factors

	1971-1994 (39 Mishaps)	1995-2002 (6 Mishaps)	TOTAL (45 Mishaps)
Mission Planning	**18**	**5**	**23**
EB101--FLIGHT PLANNING	5	4	9
EB104--FLIGHT BRIEFING	18	3	21
EB105--WEATHER ANALYSIS.	6	1	7
IB603--INFLIGHT ANALYSIS	1	3	4
IB604--INFLIGHT PLANNING	1	0	1
Crew Coordination	**25**	**5**	**30**
EB201--CREW COORDINATION	24	5	29
EB204--LEADERSHIP	2	1	3
EB205--RANK IMBALANCE	3	0	3
EB206--SUBORDINATE STYLE	4	4	8
EB404--CREW MAKEUP	1	1	2
IB801--COMPLACENCY	6	3	9
IB802--EXCESSIVE MOTIVATION	15	1	16
IB804--GET-HOME-IT IS	6	1	7
IB805--CONFIDENCE	5	0	5
IB811--OVERCONFIDENCE	13	0	13
IB812--PREOCCUPATION	8	2	10
IB9O2--CONSERVATIVE	2	0	2
IB905--INVULNERABLE	2	0	2
IB910--SUBMISSIVE	4	1	5
Task/Time Management	**26**	**6**	**32**
IB203--COG. OVERSATURATION	11	5	12
IB401--CHECKLIST ERROR	4	0	4
IB402--INADVERTENT OPS	5	0	5
IB403--NAVIGATION ERROR	6	0	6
IB404/5--WRONG SEQUENCE/SWITCH	1	0	1
IB406--WRONG TECHNIQUE	3	0	3
IB601--ACTION DELAYED	21	2	23
Situational Awareness	**32**	**5**	**37**
IA205--VISUAL ILLUSION	6	1	7
IA206--MISPERCEIVE DISTANCE	2	0	2
IA207--MISPERCEIVE SPEED	12	0	12
IB202--CHANNELIZED ATTN	24	2	26
IB205--DISTRACTION	18	2	20
IB208--INATTTENTION	15	1	16
IB209--SELECTIVE INATTENTION	5	1	6
Decision Making/Risk Assessment	**31**	**6**	**37**
IB602--WARNING IGNORED	6		6
IB605--INTENTIONAL FAILURE TO USE ACCEPTED PROCEDURE	19	1	20
IB606--RISK ASSESSMENT	5	6	11
IB608--COURSE OF ACTION SELECTED	22	4	26
IB610--FLIGHT VIOLATION	13	2	15
IB806--DISCIPLINE	4	1	5
IC102--PEER RULE VIOLATIONS	3	0	3
Communication	**14**	**4**	**18**
IC302--BODY LANGUAGE	0	0	0
IC303--COMMUNICATION HABITS	2	0	2
IC304--DISRUPTED COMMUNICATIONS	6	0	6
IC305--EXTERNAL COMMUNICATION	2	2	4
IC307--INTRACOCKPIT	4	3	7
IC308--MISINTERPRETED	10	2	12
IC309--RADIO DISCIPLING	0	0	0
IC310--VOICE TONE/INFLECTION	0	0	0

Situational awareness appears to be much less frequently cited in recent years (See Figure 5 for comparisons of mishap rates for each CRM area). This includes historically common problem areas such as channelized attention, distraction, and inattention. Strong trends in this area have not emerged in recent years.

In the area of ***decision making/risk assessment***, there has been a marked reduction in the number and rate of intentional failures to use accepted procedures (historically 49%, since 1995, 17%). On the other hand, there has been a big jump in mishap rates in the area of risk assessment, where the rate of Class A mishaps with this element rose from 13% historically to 100% of Class A mishaps since 1995. This has prompted calls for enhanced risk assessment processes in recent mishap recommendations and training.

The mishap rate for course of action remains fairly high--55% historically and 66% since 1995. Significant advances have been made in our understanding of how experts make good decisions under stress. Klein (2000) provides a superb summary of recent advances in naturalistic decision making and applies emerging concepts in a blueprint for improving how we train pilots to make better decisions. Several implications for CRM training are emerging from Navy research (Gillan, 2002) investigating decision making by novice and expert S-3 crews in a simulated flight environment.

Finally, ***communication*** is the CRM area that is cited least often as a mishap factor. Communication factors are unevenly distributed across elements (Chi Square=33.26, df=7, $p<.001$). Misinterpreted communication was historically cited relatively frequently, and still is, being cited in a third of recent mishaps. A claim made in some communication courses is that only a small percentage of the message is based on content, with much of the message being communicated through nonverbal means (e.g., body language and voice tone/inflection). It is interesting to note that none of the mishaps in our analyses cited these nonverbal elements. Intracockpit communication was cited in half the post-1995 Class A mishaps, and external (outside the cockpit) communication was cited in one-third of these recent mishaps. In all cases, the mishap narrative focuses on either a message never sent or a message not received.

CONCLUSIONS

There were two outcomes of note in our training records analysis. The first was the lack of sensitivity in quantitative student record data--instructor ratings of skill and knowledge--for diagnostic purposes. The second was the quantity and quality of instructor comments. Quantitative and qualitative analyses of these comments were combined with insights gained from CRM observations made during MOST missions to justify and then guide a major effort to improve the CRM curriculum at 58 SOW. The student behaviors described in instructor comments suggested that CRM "bottlenecks" on the flightline could be traced to a lack of specific skills and coordination processes such as repeating communication and assertiveness problems during early tactical missions. These trends led to behaviorally-based CRM training objectives. The analyses also contributed to the realization that CRM training objectives were appropriate for the hands-on portions of training as well as academic instruction.

Given this experience, instructor comments would likely be more useful than more traditional quantitative ratings when assessing the impact of explicit training interventions, such as new courseware, an enhanced instructor operator station, upgraded simulator, or reduced flight hours. In these cases, a baseline measure of student proficiency could be established through quantitative analyses of instructor comments, with a post-implementation trend plotted for statistical comparison. Instructor comment analyses can be used to extract student proficiency “snapshots over time,” track and trend problem areas for remediation, identify training missions posing greatest difficulties, flag problem students in advance, as well as look for internal weaknesses in a training program (e.g., unnecessary training sessions, chronically weak CRM behaviors).

Preserving Air Force personnel and material resources is one of two CRM program goals, the other being to maximize operational effectiveness and combat capability. Much of our earlier research focused on the latter. Mishap report analyses enabled us to focus on the first. Both, of course, have merit. We were able to identify specific, repeating trends in each of the six CRM areas to help guide course content in the next-generation CRM program at Little Rock AFB. Frequently occurring factors such as risk assessment are now being analyzed in greater detail in both the original mishap reports and in our parallel MOST simulator study to identify key behaviors.

Maurino (1999) makes an eloquent argument concerning limitations of mishap data as the sole source upon which to develop CRM instruction. We concur. In fact, CRM appears to be a multi-faceted concept, and as such, viewing it from any single vantage point is unlikely to capture its real nature. A more productive approach is to take advantage of multiple data sources. Both instructor comments and mishap reports allow training analysts to leverage the considerable efforts of others. We believe that both are efficient and powerful data sources that should be routinely used to broaden the scope of training analyses.

REFERENCES

Bruce, P.D., Killion, T.H., Rockway, M.R., & Povenmire, H.K. (1991, August). *B-52 and KC-135 mission qualification and continuation training: A review and analysis*. (AL-TR-1991-0010, AD A241 591). Williams AFB, AZ: Aircrew Training Research Division, Armstrong Laboratory.

Cronbach, L.J. (1990). *Essentials of psychological testing*. (5th edition). New York: Harper Row Publishers.

Gillan, C.A. (2002). *Aircrew adaptive decision making: a cross-case analysis*. Doctoral dissertation, University of San Diego, San Diego CA.

Hays, W.L. (1973*). Statistics for the social sciences (2nd edition*). New York: Holt, Rinehart, & Winston.

Helmreich, R.L., Merritt, A.C., and Wilhelm, J.A. (1999). The evolution of crew resource management training in commercial aviation. *The International Journal of Aviation Psychology*, 9, 19-31.

Klein, G. (2000). How can we train pilots to make better decisions? In *Aircrew training and assessment,* (O'Neil and Andrews, eds.). Mahwah, NJ: Lawrence Erlbaum Associates.

Luna, T.D. (2001). USAF Aviation Safety: FY 2000 in review. Available at http://safety.kirtland.af.mil/.

Maurino, D.E. (1999). Safety prejudices, training practices, and CRM: A midpoint perspective. *The International Journal of Aviation Psychology*, 9, 413-422.

Salas, E., Prince, C. Bowers, C.A., Stout, R.J., Oser, R.L., & Cannon-Bowers, J.A. (1999). A methodology for enhancing crew resource management training. *Human Factors*, 44, 162-171.

Silverman, D., Spiker, A., Tourville, S., & Nullmeyer, R. (1997). Team coordination and performance during combat mission training. In *Proceedings of Interservice/Industry Simulation, Education, and Training Conference, 1997*. Orlando, FL.

Spiker, V.A., Tourville, S.J., & Nullmeyer, R.T. (1999, 19 January). *MC-130P mission qualification training records and student proficiency*. (Technical Memorandum). Mesa, AZ: Aircrew Training Research Division, Armstrong Laboratory.