2011

Mishap Reduction Training for C-130J Crews

Robert Nullmeyer

Alan Spiker

Greg Deen

MISHAP REDUCTION TRAINING FOR C-130J CREWS

Robert Nullmeyer
711th Human Performance wing
Mesa AZ

Alan Spiker
Anacapa Sciences
Santa Barbara CA

Greg Deen
Lockheed Martin Simulation, Training, and Support
Little Rock AFB AR

The C-130J is an advanced automation aircraft flown with a smaller crew than its tactical airlift predecessors. It is seeing increased action in theater. The Air Mobility Command sponsored a multi-prong project to improve C-130J aircrew training for operations in high threat environments: (1) analyze C-130J and related platform (C-130 E/H, C-17) mishap reports, (2) survey C-130J crew real world experiences regarding threats to safety, and (3) observe C-130J crews as they plan and execute a challenging, tactically relevant simulator scenario. The Air Force C-130J safety record is enviable—no crew-caused Class A mishaps and low rates across all mishap categories relative to other airlift platforms. Human factors frequently cited in Class B and C reports included checklist errors, distraction, task prioritization, and decision making. Incident and mishap reports both frequently mention problems arising from events external to the crew. In the real-world experience survey, several crews described events that closely paralleled events in the simulator scenario. During the challenging simulator scenario, crews generally accomplished the mission but quality of performance varied considerably. Lower performing crews often struggled with risk management during planning and mission evaluation during execution. They also tended to under-utilize their loadmasters and had difficulty choosing the most appropriate levels of automation during mission execution. Recommendations included an increased emphasis on threat and error management during training and addressing specific skills at particular points in the syllabus.

A longstanding and central role of human error in aviation mishaps is well documented. For example, Helmreich and Fouchee (1993) reported that, from 1959 to 1989, flight crew actions were causal in more than 70% of worldwide air carrier accidents involving aircraft damaged beyond repair. More recently, researchers are reporting diminishing proportions of air carrier mishaps world-wide being attributed to operator error. Baker, Qaing, Rebok, and Li (2008) reported a statistically significant drop in air carrier mishaps involving human error, from 42% in the 1980s to 25% in 1998-2002. Dismukes, Berman, and Loukopoulos (2007) reported that rates of “crew caused” accidents dropped by 50% between 1978-1990 and 1991-2001. While mishaps attributed to crew error may be declining, they still represent a large portion of the total count, so the nature of that error remains an obvious issue. The National Transportation Safety Board (NTSB) reported that 68% of crew-caused accidents from 1989-1990 involved tactical decisions, especially the failure to execute a go-around given an unstable approach, and 84% involved inadequate monitoring/ challenging (NTSB, 1994). Dismukes and his colleagues analyzed more recent (1991-2001) NTSB reports and found 19 US air carrier accidents where crew error played a central role. Descriptions were provided for each of these events, and accident statistics were compared with those from the earlier NTSB analysis. They found that crews faced the same basic challenges in more recent years--74% of recent crew-related accident reports cited tactical decisions and 68% cited monitoring/challenging errors.

By the late 1990s, it became increasingly recognized that errors were only part of the safety picture and that a broader understanding of all threats to safety was essential to develop effective responses, and several airlines moved toward a threat and error management (TEM) model of safety interventions (Helmreich, Klinect, & Wilhelm, 2001). TEM is an organizational response to safety that includes analyses of operational data from line oriented safety audits, flight data recorders, and crew self reports of hazardous situations. The goal is to understand and manage threats to safe operations. From these data sources, several solutions may emerge, including changes in operating procedures or equipment and tailored training interventions. This TEM approach in commercial aviation training is recognized as a “best practice” by the International Civil Aviation Organization, the International Air Transport Association, the National Air Transport Association, and the U.S. Federal Aviation Administration.

In contrast to trends in airline safety, a recent Air Force review of accidents through 2009 revealed that percentages attributed to human factors remained relatively constant at about 70% (Heupel, Gardetto, Dopslaf, Hughes, Williams, & Johnson, 2010). A recent review of mishap rates across *all* Unites States military services revealed little systematic improvement in mishap rates over the past 20 years (Unites States Coast Guard Safety Center, 2009). O'Connor, Hahn, & Nullmeyer (2010) recently reviewed military Crew Resource Management (CRM) training across both the United States armed services and other allied countries, and found no military training programs today that pursue TEM as aggressively as has been the case in commercial aviation.

Air Mobility Command sponsored a series of efforts to better understand threats to safety for C-130J crews and how these crews responded to those threats: (1) review relevant safety data; (2) survey mission qualified crews to solicit accounts of "threats" to flight safety, some type of crew error, and/or a successful mitigation strategy or technique; and (3) observe mission qualified crews perform a tactically demanding capstone simulator scenario during annual refresher training. The overall objective was to provide actionable recommendations for updating C-130J training to better reflect TEM concepts. The focus in this paper is on the first and third of these elements.

Results

Accident and Incident Trends.

The Accident Investigation Board (AIB) homepage (http://usaf.aib.law.af.mil) summarizes AIB reports from Class A accidents (fatality or permanent disability, loss of aircraft, or more than $2 million damage). The Air Force Safety Center (AFSC) home page (http://afsafety.af.mil/) provides considerable summary mishap statistical information, including hours flown and mishap frequencies, by aircraft type and year. Flying hours per year are essential for translating frequencies into mishap rates. AFSC also maintains an electronic database of accident reports, as well as high accident potential and hazardous air traffic self-reports. Data from all of these sources were used to review Air Force airlift mishap trends.

C-130 Class A mishap rates across all models fell consistently from 1970 through 2000, with .73 mishaps per 100,000 flying hours in the 1970's, .56 in the 1980's, and .26 in the 1990's. The rate increased slightly from 2000-2009, to .35 per 100,000 flying hours, and that pattern of slight increases in Class A mishap rates was common throughout the Department of Defense. In response to rising mishap rates, all Services began a series of safety initiatives in 2005. C-130 class A mishap rates dropped from 2.9 in 2001-2005 to 1.6 from 2006-2010, suggesting that the safety initiatives may be proving successful. C-130J flight operations began in the late 1990's. The C-130J safety record compared to earlier C-130 models of is enviable – one Class A ground mishap in 2008. There have been approximately 150 C-130 Class B mishaps in the past decade. Six of these involved a J model C-130, and of these, only one involved crew factors. Two conclusions were drawn from these trends. First, the drop across the past two decades that has been documented in air carrier human factors mishaps was not as pronounced in Air Force accident data. Second, the C-130J community has had an enviable safety record to date.

We broadened the scope of mishap analyses to include C-130E and H Class A mishaps given the shared tactical airlift mission, and to include C-17 mishaps given similarities in cockpit display technology. Mishaps occurred from 1999 through 2009. In both of these other aircraft, numerous mishaps involved human error. There were 13 C-130 E and H Class A mishaps, of which eleven had notable crew factors. Of 22 C-17 Class A mishaps, nine involved notable crew factors. The top human factors in C-17, legacy C-130, and C-130 J mishaps are shown in Table 1. For the C-130J, the human factors came from one Class B mishap and eight Class C mishaps.

The factors in Table 1 come from reports where human factors analyses were based on the Department of Defense Human Factors Analysis and Classification System, or DoD HFACS. DoD HFACS is well documented in multiple places, including the Navy and Coast Guard safety center web sites. Several patterns emerged from our analysis. Inattention was frequently cited in both aircraft with highly automated cockpits (C-17 and C-130J) but not legacy C-130 mishaps. A similar pattern emerged with the highly related factors of cognitive task oversaturation and task misprioritization. Both refer to challenges with managing multiple tasks simultaneously. Risk assessment and decision making are also highly related. HFACS distinguishes between risk assessment and decisions during planning from those behaviors during mission execution. All three platforms cite this skill area during mission execution. Violating stated rules shows up in all three lists. Procedural guidance/publications appears in C-17 and C-130 accidents. A common theme when citing this factor is the difficulty keeping crews current given rapidly changing tactics, techniques, and procedures. While this was not among the top factors in C-130J accident reports, it has emerged as a common factor across the entire Air Force (Air force Safety Center, 2009).

Table 1. The Most Frequently Cited Mishap Human Factors in Three USAF Aircraft Types

C-17 (Class A)	C-130 E/H (Class A)	C-130J (Class B & C)
Risk assessment during operations	**Procedural guidance/publications**	**Distraction**
Procedural guidance/publications	**Risk assessment during operations**	**Task misprioritization**
Inattention	**Mission planning**	**Decision making during operations**
Cognitive task oversaturation	**Planned inappropriate operations**	**Checklist error**
Channelized attention	**Violation - lack of discipline**	**Violation - lack of discipline**
Fatigue physiological mental	**Channelized attention**	**Inattention**
Planned inappropriate operations	**Local training**	**Complacency**
Overcontrol/undercontrol	**Necessary action delayed**	**Misperception of conditions**
Violation - lack of discipline	**Miscommunication**	**Breakdown in visual scan**

Air Force crews are increasingly documenting near misses in flight or on the ground through Hazardous Air Traffic Reports or HATRs. Over 700 HATRs have been submitted involving C-130s of all types from 1999 through 2009. Of these, 20 involved C-130J aircraft. HATRs tend to be short (about 2 pages) summaries of what transpired. Air Traffic Control errors were cited in half of these HATRS, and errors by other pilots were cited in most others. C-130J crew errors were contributing factors in only two of these incidents, further supporting the need to address externally generated threats to safety, rather than focus solely on C-130J in-cockpit error.

Crew Performance in a Simulator-Based Tactical Scenario.

Participants: Twenty mission qualified C-130J crews were observed as they planned a realistic, tactically complex mission; executed that mission in a full crew, high fidelity weapon system trainer; and debriefed the mission. A C-130J crew consists of an aircraft commander (AC), a co-pilot (CP), and a loadmaster (LM).

The Scenario: Annual refresher training for C-130J crews builds up to an exercise where mission qualified crews plan a challenging tactical airlift mission and then execute it in a C-130J Weapon System Trainer, a full crew, high fidelity simulator. An Afghanistan scenario was used in this study. It was specifically designed to interject a number of "threats" or problems to the crews, beginning in planning and continuing throughout the mission. For example, in planning, a Block 5.4 aircraft was specified. This has weight, fuel, and takeoff performance implications that should be considered during planning. Ground threats enroute and around the landing zone were briefed, and the landing zone runway sloped steeply on one end. During execution, challenges included fairly stringent altitude restrictions in the initial climb-out, a marked shift in wind direction affecting an already demanding approach to the landing zone, selected equipment failure, lack of unloading equipment where cargo was being delivered, a cargo fire, and loss of an engine that necessitated a divert decision. Many similar challenges were mentioned by crews in our real-world TEM survey.

Data Collection. Crew performance was independently rated by two observers who used structured data collection forms that were tailored to the scenario being observed. All ratings were based on five-point behaviorally anchored scales, ranging from 1= poor to 5 = exceptional. One observer rated crew CRM processes during three phases of the exercise (planning, execution, and debrief). A second observer rated crew performance. CRM ratings addressed the six Air Force Instruction 11-290 content areas of mission evaluation, task management, situation awareness, crew coordination, communication, and risk management/decision making. Crew performance during planning was rated as seven separately graded items contingency considerations, takeoff and landing data (TOLD), and decision quality, among others. During execution, 14 items were rated including use of automation, checklists, time control, aircraft handling, response to emergencies. The two observers rated their respective areas independently, and did not compare scores or try to reach consensus.

Data Analysis. An initial analysis of the C-130J simulator refresher study data addressed the relationships between ratings that were assigned by the observer for crew process and the observer for mission performance. Subsequent analyses addressed (1) how well the 20 crews handled the “threats” that were administered to them during the session, (2) observer comments that characterize the process behaviors of the most and least successful crews, and (3) the influence of demographic background on CRM behaviors and mission performance.

Overall Ratings. One major question to be answered is the degree to which CRM crew processes predicted mission performance. Figure 1 provides a graphic of the mission execution data from the 20 crews in which the x-axis corresponds to the average CRM process rating received with the y-axis indicating mission performance. Since some crews have identical ratings for CRM process and mission performance, we have indicated multiple crew presence by making the dot bigger with a circle around it. Also, each crew in that position is identified by a crew number above and to the right of the dot. A sizeable positive correlation was observed between crew process and mission performance ($r = .60$, $t = 3.17$, $df = 18$, $p < .01$). In keeping with this relationship, the majority of the crew plots should be located in quadrants I and III. Indeed, from Figure 1 we can see that only one crew, Crew #12, is found in Quadrant II, which is indicative of higher mission performance (average = 4) than would be predicted based on their CRM process rating (average = 2). No crews are found in Quadrant IV (lower performance, higher process). Moreover, two other crews can be said to have a “mismatch” between CRM process and mission performance, as defined by a difference in rating greater than 1 unit. These are Crew #7 (process = 3, performance = 1) and Crew #10 (process = 4.5, performance = 3). The remaining 17 crews are in line with a consistency in ratings between CRM process and mission performance. In fact, 11 of our 20 crews have identical ratings for CRM and performance, either (2, 2) (Crew #4, 11, 13, 14, 20), (3, 3) (Crew #3, 19), (4, 4) (Crew #6, 8, 15), or (5, 5) (Crew #17). For the most part, then, crews that have better CRM process behaviors had better mission performance.

Characteristics of “Strong” and “Weak” Crews. We can use this consistency between CRM and performance to unequivocally identify the strongest and weakest refresher crews during this scenario. Specifically, we labeled the crews whose process-performance plots fall high in Quadrant I as our strongest crews. Looking at Figure 1, we can see by this definition that our four “strong” crews are: #6, 8, 15, and 17. Similarly, we looked to crew plots in Quadrant IV for our “weak” crews. Here, we had five such crews: #4, 11, 13, 14, and 20. In subsequent analyses, we examined the observer protocols for the strongest crews to identify best practice CRM behaviors and corresponding mission performance elements that are representative of what the most successful crews do. In like fashion, we extracted typical crew interaction breakdowns, errors, and associated problems with our weakest crews to generate a list of “avoid these” behaviors. Several highlight of that analysis follow:

- Higher quality mission planning, both in terms of time spent and activities performed, resulted in superior performance. Strong crews tended to plan for the most likely contingencies, particularly with regard to go-arounds at the LZ. Indeed, we observed that thorough LZ study was a key performance-determinant in the study.
- Without exception, the most successful crews in the simulator study treated and utilized the LM as an integral part of the crew throughout the mission. Indeed, we found that effective utilization of the LM was a notable predictor of mission performance in its own right. Thus, it is quite clear that the C-130J, despite its advanced automation, is a true 3-person cockpit (at least during parts of the mission) rather than a 2- or 2.5-person cockpit.
- The ability of crews to program and re-program take-off and landing data (TOLD) during the mission was a major determinant of success. In particular, crews who accepted the TOLD data given them in planning without verification suffered high workload later as the threats began to mount up. Moreover, failure to check data and verify calculations caused some serious problems during LZ approach and landing, resulting in a less successful mission. Strong crews actively calculated and verified as much TOLD data as possible while on the ground.
- Effective set-up and use of automation was a major characteristic of successful crews and its absence was typically a reliable predictor of poor performance. From our analysis of the simulator study data, it became clear that automation use is a continuum in which either extreme – over-reliance on automation to the exclusion of pilotage skills or under-utilization such that the pilot spends the majority of his/her time hand flying – will result in degraded performance. Indeed, several crews where the aircraft commander rarely used the autopilot were extremely task-saturated, resulting in poor mission performance.

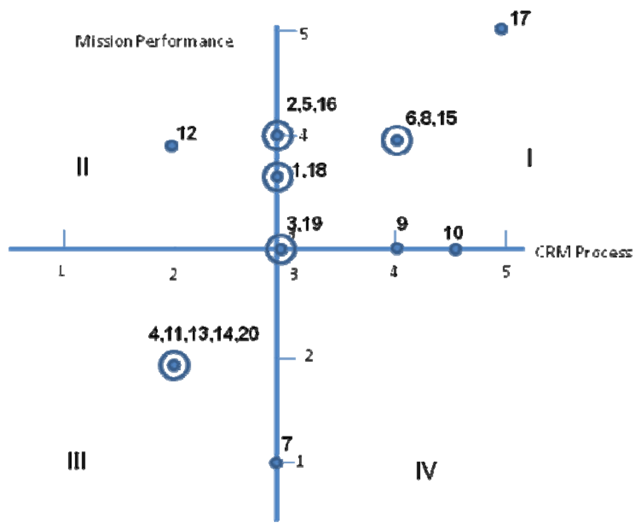

Figure 1. Crew plots of average crew process and mission performance ratings.

- Fortunately, most crew members took the mission seriously, not just as a simulator session, and did their best throughout. However, there were several crews who treated the mission as a typical CONUS operation despite the clear difficulties described during planning. This attitude, coupled with a lack of clear leadership and a closed or partly-closed environment for crew member inputs, creates a bad mix for conducting a mission and, not surprisingly, crews with such an attitude scored at the bottom on mission performance.

Crew Demographics and Mission Performance. Average C-130 flying hours for both C-130J and legacy C-130 are shown in Table 2 for members of both stronger and weaker crews. Comparing the top two rows, we see that there are some fairly sizeable differences between the two groups in average experience. However, because the within-group variances are also fairly large, we elected to perform independent t-tests to determine if any of these differences reached statistical significance. The bottom row depicts the probability of a significant difference given a t-test with seven degrees of freedom. As can be seen, the only significant strong vs weak crew difference is for the LM legacy C-130 hours. Specifically, crews whose LM's had a large number of legacy C-130 flying hours were more likely to part of a crew that was classified "weak" rather than "strong." Crews where the LM had extensive legacy C-130 experience – where the LM was viewed as a backend crewmember only – were at a disadvantage.

Table 2. Average Flight Experience for "Strong" and "Weak" Study Crews.

Crew Classification	**Average AC C-130J Hours**	**Average AC Other C-130 Hours**	**Average CP C-130J Hours**	**Average CP Other C-130 Hours**	**Average LM C-130J Hours**	**Average LM Other C-130 Hours**	**Average Total J Hours**
Strong 1138		1556	968	745	1326	417	3431
Weak 930		910	896	1410	1660	1890	3486
p-value (based on indep t-test)	.63 .	46	.86	.43	.30	**.03**	.93

Implications for Training

The C-130 Hercules aircraft has been in production for over 50 years, serves untold missions, and is utilized by numerous services around the world. The advancements in aircraft technology were minor when compared to the introduction of the "J" version. The "classic" C-130 typically required a crew of four in the cockpit, and, for the USAF, two LMs in the cargo compartment. Appropriate flying speeds, fuel endurance, route headings and weight and balance of cargo were calculated manually using charts and slide-rule technology. The navigators used sextants for overwater route navigation into the ‘80s, slowly replaced by INS and GPS systems. The C-130J reduced the flight crew to two pilots and one loadmaster by taking advantage of advances in computers and avionics.

The military mission, however, became more complex. It only stood to reason that the skill-sets taught to the legacy aircrews needed to change, and perhaps since a flight engineer and navigator are no longer in the cockpit does not mean their activities are no longer required. On the contrary, the tasks of the past exist; those and many more are now within the computers on board. The LMs no longer crank a bicycle chain to unlock cargo or physically remove a connecting link to fully lower the aft ramp; remote electric switches now perform that task. Training aircrews to understand and interface with the computers that operation the aircraft systems will be the major challenges to convert legacy crews to the automated operating system. This report confirmed that the pilot’s prowess with the automation is a key requirement for training. Workload management, even with the aid of automated systems, can still overwhelm a 2-pilot crew. It would seem advisable, then, to encourage the 3-person cockpit concept during refresher training, particularly for crews where extensive legacy C-130 experience is present. When considering the emerging objective of TEM, adaptability will become a key skill for the crews. Tactical airlift, by its nature, includes short-notice changes to the planned mission.

This report clearly indicates the need for increased training in the understanding and use of automation to ensure pilots are fluent in the operating system of the flight management computers, and the 3 crewmembers perform with naturalistic synergy. Loadmasters are no longer just “cargo-guys in the back” but now must play a key role in supporting cockpit activities, particularly in times of high taskloads and changes in mission requirements. While the C-130J aircraft is a significant advancement in aircraft technology, it demands equally significant changes in CRM skills. These changes call for increased mission performance training, as a crew, in realistic simulator scenarios. To address this, the C-130J training system for the USAF will introduce a thoroughly revised training curricula beginning FY 2012. This new plan will incorporate a new part-task trainer that emulates the automated operations of the flight deck. These training aids will concentrate on automation-management skills that are currently taught during simulator lessons. Loadmaster fuselage trainers will also have these training aids to allow the loadmaster students opportunity to acquire proficiency in the cockpit tasks needed to support the pilots prior to their joining the pilots in the flight simulators. This shift in training will allow a reduction of simulator events within the training program, and allow the simulator missions to focus more on CRM and mission-accomplishment skills by the entire crew, for which the full crew, full-motion simulator is better suited.

References

Baker, S.P., Qiang, Y., Rebok, G.W., and Li, G. (2008). Pilot error in air carrier mishaps: longitudinal trends among 558 reports, 1982-2002. *Aviation, Space, and Environmental Medicine*, 79(1), 2-6.

Dismukes, R. K., Berman, B. A., & Loukopoulos, L. D. (2007). *The limits of expertise: Rethinking pilot error and the causes of airline accidents*. Aldershot, UK: Ashgate Publishing Co.

Helmreich, R.L. & Foushee, H.C. (1993). Why crew resource management? Empirical and theoretical bases of human factors training in aviation. In E.L. Wiener, B.G. Kanki, and R.L. Helmreich (Eds.), *Cockpit resource management*, San Diego, CA: Academic Press.

Helmreich, R.L., Klinect, J.R., & Wilhelm, J.A. (2001). System safety and threat and error management: the line operations safety audit (LOSA). In Proceedings of the Eleventh International Symposium on Aviation Psychology (pp. 1-6). Columbus OH: The Ohio State University.

Heupel, K., Gardetto, P., Dopslaf, E., Hughes, T., Wiliams, S. & Johnson, B. (2010). USAF aviation safety: FY2009 in review. Presented at the 2010 annual meeting of the Aerospace Medical Association, Phoenix AZ.

Loukopoulos, L.D., Dismukes, R.K., & Barshi, I. (2001). Cockpit interruptions and distractions: a line observation study. In *Proceedings of the Eleventh International Symposium on Aviation Psychology*. Columbus OH: The Ohio State University.

National Transportation Safety Board [NTSB] (1994). *A Review of Flightcrew-involved Major Accidents of US crriers, 1978 through 1990.* (Report No. PB94-917001, NTSB/SS-94/01.

U.S. Coast Guard (2008). *FY07aviation safety report,* available at http://www.uscg.mil/safety.