2011

Human Performance Training: Successes and Failures in Civil Aviation

Alan E. Diehl Ph.D.

HUMAN PERFORMANCE TRAINING: SUCCESSES AND FAILURES IN CIVIL AVIATION

Alan E. Diehl, Ph.D., ATP, CPE
Retired NTSB, FAA & USAF Air Safety Investigator
Albuquerque, New Mexico

For over twenty years, two similar types of training have attempted to reduce aviator caused accidents. Crew Resource Management was widely adopted by U.S. airlines, and has generally been credited with helping to dramatically reduce their accident rates. Interestingly, CRM was embraced with little scientific evidence that it could actually reduce operational errors. In contrast, Aeronautical Decision Making, aimed at general aviation users, underwent a series of double-blind experiments before being adopted. And, although some users suggested ADM training significantly reduced their accidents, it was never fully implemented. This may explain the limited improvement in general aviation accident rates, when compared with the airline rates for the last two decades.

"Human error" has historically been associated with approximately three quarters of all aviation accidents, and various studies have indicated aviator decision making failures were associated with the majority of fatal civilian crashes, (e.g., Diehl, 1992). Two major training efforts were launched in the 1980s to deal with these problems. These similar programs, Crew Resource Management and Aeronautical Decision Making, were proffered as cost-effective and quick methods of reducing such accidents.

Crew Resource Management

While serving as a National Transportation Safety Board investigator, I drafted the first recommendation calling for the implementation of Crew Resource Management training by this nation's airlines. This occurred after a 1978 United Airlines DC-8 crash. The crew detected a landing gear unsafe light, and entered a holding pattern, only to run out of fuel 66 minutes later over the suburbs of Portland, Oregon. The captain, fixating on the gear light problem, had ignored repeated queries from the other two crew members about their fuel status, (NTSB, 1979).

Per the NTSB recommendation, United instituted CRM training in 1981. And within several years most U.S. major carriers followed suit. Although, the Federal Aviation Administration did not formally require CRM for all FAR Part 121 carriers until 1995.

NTSB investigators must carefully justify proposed recommendations. So I examined the limited evidence that CRM training would work. I reviewed the research conducted earlier under the direction of Dr. John Lauber at the NASA Ames Research Center, including the simulator studies done by Dr. H. P. Ruffell-Smith. In addition, contemporary research on the recognition of "subtle incapacitation" done at United Airlines as well as Line Oriented Flight Training being conducted by Northwest Airlines, looked promising.

As an adjunct professor of management, I also recognized the similarity of CRM to another type of training used in industrial settings. Total Quality Management empowered workers to speak-up and actively participate in decision making about products and production processes. TQM had been successfully utilized for decades and was credited with helping the

successful renaissance in Japanese industry after World War II. Ironically, it was the brainchild of an American, Dr. Edwards Deming.

CRM-related errors also seemed to be factors in several well-known crashes. These accidents included the 1972 Eastern Airlines L-1011 crash into the Everglades and the tragic collision of the KLM and Pan-American B-747s on the runway at Tenerife in 1977. Such catastrophes suggested that issues like loss of situational awareness, communication breakdowns, as well as individual and collective judgment errors needed to be addressed.

Thus, adopting CRM training seemed logical, even without any statistical data or controlled experiments to "prove" it worked. In any case, I felt a new type of training to manage resources on the flight deck was an idea whose time had come, and the NTSB members, as well as United's management and the union apparently agreed.

The FAA authorized an18-month evaluation of the prototype CRM program. The Agency even assisted United to defer some of the costs of the program's development, and the two-day training sessions for its flight crews. The FAA did this by granting the airline an "exemption" allowing its captains to take line checks, instead of going to simulator training every six months. Thus, captains only received simulator training once a year. I was aware of this arrangement, because I had transferred from the NTSB to become the FAA headquarters program scientist for human performance, in 1981.

However, some people legitimately questioned whether this type of training *really* worked. The term "psycho-babble" was used by some FAA bureaucrats. But I retorted that CRM training provided a host of pragmatic techniques to reduce the probability and the consequences of errors -- what others would later call "threat and error management."

Other airlines soon approached the FAA requesting exemptions for their proposed CRM programs, but not all of these programs were well designed. This was a legitimate concern for the Agency. So, in 1983, I hired Dr. Richard Jensen to examine cutting-edge CRM training programs in use around the globe, and to show the FAA how best to certify such programs and airmen enrolled therein.

I was also delighted to hear that NASA was also developing instruments like the Cockpit Management Attitude Questionnaire to evaluate the views of airmen about CRM. However, such measures could not answer the question of whether this training reduced the number of errors during "normal" much less emergency operations, (Wiener, Kanki & Helmreich, 1993).

But perhaps the best evidence that CRM actually worked came from a couple of accidents experienced by United Airlines crews after being exposed to the concepts for several years. In August 1989 a United DC-10's center engine exploded in-flight, severing hydraulic lines and eliminating all aerodynamic controls. The only way to control the jet was through differential power of the two remaining engines, (NTSB, 1990a).

The DC-10 crew led by Captain Al Haynes allowed another off-duty captain handle the thrust levers, and they were able to successful crash land at the Sioux City airport. This widely reported feat of airmanship saved the lives of the majority of the passengers aboard the ill-fated craft. Captain Haynes has repeatedly stated the use of CRM was essential in their success that fateful afternoon, (Haynes, 1991).

Several months earlier another United crew faced an almost as demanding situation. Their 747 had departed Honolulu, but as they passed through 22,000 feet, a forward cargo door failed, ripping away part of the aircraft's fuselage. The explosive decompression ejected nine passengers and baggage, and damaged both starboard engines, along with the wing's leading edge devices, (NTSB, 1990b).

Using various CRM techniques they managed to safely return to Hawaii without the further loss of life. For instance, the second officer soon recognized they were far too heavy for the rapidly up-coming two-engine-out emergency landing. But, he just could not get the bewildered captain's attention, so he quickly decided to dump fuel himself. I appreciated the irony of this situation, because in my 1978 United DC-8 crash, the second officer could not get the captain's attention about their *low* fuel state. Certainly, CRM techniques played a role in saving this 747 and its passengers.

Obviously, trying to make precise statistical statements about extremely infrequent events is always a problem. It has been noted elsewhere that because airline accidents are so infrequent it is difficult to establish the precise role that CRM has played in the reduction in these accidents, (Helmreich, Merritt & Wilhelm, 1999).

Furthermore, many things besides the use of CRM have changed in the aviation system in recent decades. These changes included: the advent of highly automated airliners, more effective air traffic management procedures, installation of newer types of cockpit warning devices, all of which probably helped enhance airline safety.

Incidentally, the USAF Military Airlift Command began using CRM in 1985. This change was ordered by their commander and took place in just three-months. While military transports faced different flying conditions, often with less experienced crews, it was interesting to examine the statistical changes that occurred after this program began.

I computed there was a 51% reduction in mishaps during the five-year period after the program's implementation, when compared with the previous five-year period. This was statistically significant at the .05 level. And unlike the protracted airline implementation period, little else changed for this command during that decade-long time span. Thus, this military evidence suggested CRM was a factor in the improved airline safety, (Diehl, 1992).

Aeronautical Decision Making

The other major type of human performance training focused on general aviation pilots. It was originally called "judgment training" and the initial research was led by Dr. Richard Jensen under FAA contract in the mid-1970s. A major conclusion of his landmark study was that 52% of U.S. fatal general aviation accidents were caused by pilot judgment errors. Jensen also recommended that prototype training materials be developed and their effectiveness be measured, (Jensen & Benel, 1977).

The "judgment training" label continued through the development of the early manuals and other materials. This work was done at Embry Riddle Aeronautical University, under the leadership of Dr. Jerry Berlin. I was soon asked by FAA officials to assume the technical direction of the program as well as the evaluation of the prototype materials. The FAA wanted to insure such training was effective before proceeding further.

Judgment training materials taught many of the same concepts used by airline CRM programs, but the information and case studies were drawn from general aviation situations and single pilot operations. Subject pilots learned about the effects that several types of hazardous attitudes have on performance, as well as concepts such as stress management, risk management, and attention management.

And they were also given various "rules and tools" to help improve their decision-making. For example, the "I'M SAFE" mnemonic reminded subjects that before flying they needed to insure they were free of the following conditions: Illness, Medications, Stress, Alcohol, Fatigue, and to ensure they had Eaten properly.

A major challenge in evaluating this training was whether or not pilots would use the concepts in "normal" situations. Determining if those pilots who had received judgment training performed differently than control group subjects was a requirement. It was decided to measure the decision-making of all pilots during short cross-country flights when they did not know their judgment was being carefully observed.

After receiving the prototype judgment training materials, Embry Riddle subject pilots made 17% fewer decisional errors than student pilots in the control group who had not received this training. These results, while preliminary, suggested pilot judgment could be improved through training, (Berlin, et. al., 1982).

Various organizations around the world became interested in the topic. For instance, Gary Livack and his employer the General Aviation Manufacturers Association offered to help develop cutting-edge audiovisual materials along with improved training manuals.

Dr. Georgette Bush of Transport Canada led the effort there, and conducted two additional studies. One study examined private pilots and another involved pilots training for their commercial licenses. Both studies used the new and improved training manuals, and the commercial students also were exposed to more sophisticated audiovisual materials. The two double-blind studies in Canada indicated private pilots receiving judgment training averaged 9% fewer errors, while commercial pilots (who got a more comprehensive program) averaged 40% fewer errors, when compared to those pilots assigned to the respective control groups, (Bush and Diehl, 1983).

An additional study of private pilot judgment was undertaken for the Australian government by Dr. Ross Telfer. While another study, sponsored by the USAF and directed by Dr. Tom Connelly, used instrument-rated ROTC cadets. Here pilots faced simulated emergencies in GAT-1 instrument training devices.

In all five experiments, pilots receiving judgment training outperformed their contemporaries at statistically significant levels. The measured improvements averaged from a low of 8% fewer errors up to 46% fewer errors in the case of the instrument rated pilots study. This wide variation in the amounts of experimental group improvement was attributed to: the differences in the quality of the training manuals, the sophistication of other training media, and the emphasis placed on this training by instructors, (Diehl, 1992).

The FAA and the Aircraft Owners and Pilots Association wondered what would be the effect of making judgment training available, but not mandatory. So I made another study of private pilots at ten fixed base operators in the U.S. This time the experimental subjects were

simply given the training manuals without any encouragement or instructions on how to use them.

The FAA was also evaluating new sectional chart designs at the time, which provided an opportunity to observe the "normal" behavior of both experimental and control group subjects. Incidentally, the observers posed as cartographers, while unobtrusively recording the in flight judgment performance of subjects.

The results of this well-controlled double-blind study produced a 10% reduction in errors for the experiment pilots who were given the manuals. While statistically significant, the results also suggested the training would be much more effective if the subjects believed they were going to be evaluated on the materials during actual FAA flight tests, (Diehl, 1992).

And it was soon concluded the "judgment" was not the best label for such training. This was because many student pilots were older, highly successful, professional people who did not think their youngish instructors could improve their judgment. Hence, I decided to rebrand the training as Aeronautical Decision Making. And a series of ADM training manuals was developed for various categories of pilots, such as Students and Private Pilots, Instrument Pilots, Air Taxi Pilots, Helicopter Pilots, etc.

Shortly after these training manuals were completed in 1987, I transferred from FAA to the USAF, and several other key personnel associated ADM development also went on to other assignments. The result was this training was never fully implemented by the FAA. In fact, it was not until 1997 that it became "theoretically" mandatory training.

Interestingly, some people in the helicopter community immediately embraced this training. Bell Helicopters Inc. and Petroleum Helicopters Inc. vigorously pursued ADM training for the pilots flying their equipment. These two organizations soon reported dramatic drops in their respective accidents rates. Bell saw a 48% drop in their U.S. Jetrangers accidents, while PHI experienced a 54% drop in their accidents, (Fox, 1991; Adams & Diehl, 1988).

Unfortunately, unlike Bell and PHI, the broader general aviation community never thoroughly applied ADM concepts. Ironically, a 1999 study by former FAA inspector, Doug Hawley, confirmed this fact. He determined that Certified Flight Instructors spend very little time teaching ADM. In fact, of the instructors he surveyed, 33% had never heard of ADM, only 13% knew it was actually mandatory, and a mere 3% could explain how to obtain the FAA materials needed to teach these concepts, (Hawley, 1999).

Conclusions

The widely applied CRM training programs appear to have helped dramatically reduce airline accidents in recent decades. Furthermore, the failure to give general aviation pilots ADM training may well have contributed to the relatively paltry reduction in their accidents.

The U.S. general aviation accident data stands in sharp contrast with the progress made by this nation's airlines over the past 20 years. In fact, the latest data available from the NTSB (for 1990 to 2009) indicates the accident rate for all fatal FAR Part 121 airlines decreased by 78%, while the rate for all fatal general aviation accidents declined by only 15 %, (NTSB, 2010). While annual fatal accident rates, especially for airlines are often volatile, other metrics tell a similar story.

Concerns about drawing firm conclusions on the safety of such different types of flying are certainly relevant. However, the large differences in the *changes* in their respective accident rates during this twenty-year period suggests further study of the roles played by CRM and ADM training is warranted.

References

Adams, R.J. & Diehl, A.E. (1988). The Human Factor in Helicopter Safety. *Vertiflight,* July/August, 34, No. 4.

Berlin, J.I., Gruber, J.M., Holmes, C.W., Jensen, P.K., Lau, J.R., Mills, J.W. & O'Kane J.M. (1982). Pilot Judgment Training and Evaluation. DOT/FAA Report CT-82-56, Vol I, II & III, Washington, DC.

Diehl, A.E. (1992). The Effectiveness of Civil and Military Cockpit Management Training Programs. *Flight Safety Foundation, Proceedings of the 45th International Air Safety Seminar*, Long Beach, CA.

Fox, R.G. (1991). Identifying the Hazard Is Only the Beginning. *Hazard Prevention*, 27, No. 3.

Jensen, R.S. & Benel, R.A. (1977). Judgment Evaluation and Instruction in Civil Pilot Training. DOT/FAA Report RD-78-24, Washington, DC.

Hawley, D. (1999). The Implementation of Aeronautical Decision-Making in General Aviation Training Environments, Unpublished M.A.S. Thesis, Embry Riddle Aeronautical University, Albuquerque, NM.

Haynes, A. (1991). The Crash of United Flight 232, Presentation to the NASA Dryden Flight Research Facility, May 24, Edwards, AFB, CA.

Helmreich, R.L., Merritt. A.C. & Wilhelm, J.A. (1999). The Evolution of Crew Resource Management Training in Commercial Aviation. *International Journal of Aviation Psychology*, 9(1), 19-32.

NTSB (1979). Aircraft Accident Report, United Airlines, DC-8, Report No AAR-79-7, Washington, DC.

NTSB (1990a). Aircraft Accident Report, United Airlines, DC-10, Report No AAR-90-6, Washington, DC.

NTSB (1990b). Aircraft Accident Report, United Airlines, B-747, Report No AAR-90-1, Washington, DC.

NTSB (2010). Annual Review of Aircraft Accident Data: Accidents, Fatalities, and Rates, 1990 - 2009, 14 CFR 121, (table 5), and Accidents, Fatalities, and Rates, 1990 - 2009, U.S. General Aviation (table 10), Washington, DC.

Weiner, E.L., Kanki, B.G. & Helmreich, R.L. (1993). *Cockpit Resource Management,* Academic Press, San Diego, CA.