A Programme for Monitoring Fatigue in Long-Haul Commercial Operations

Dr David Powell, Dr Keith Petrie and Mark Norrie
Air New Zealand International Flight Operations

Flight Safety Foundation 51st Annual International Air Safety Seminar
16-19 November 1998

Debate around Flight Crew duty times internationally has at times been influenced by political and industrial concerns rather than just science and safety, a fact which has impeded progress in this area. This paper describes a scientifically based system currently in use operationally to assist with managing flight crew fatigue in an international airline. Air New Zealand is a small carrier with some 40 aircraft excluding the "Air New Zealand Link" feeder services. The International fleet is only 26 aircraft – 6 Boeing 747-400, 5 Boeing 747-200, 12 Boeing 767 and 3 Boeing 737-300. The route structure is characterised by long sectors because of New Zealand's geographical isolation. There are many night departures to achieve the desired arrival times at Northern Hemisphere destinations. Since the airline has no foreign basing of pilots, but operates right through to Europe, pilots are relatively frequently required to position as passengers. Our 747-400 pilots commonly only go to work twice in a month, but it can be for up to 12-13 days. These factors conspire to make circadian disruption and fatigue prominent features amongst our pilot group.

Air New Zealand was one of the first airlines in the world to introduce a policy for controlled rest on the flight deck (cockpit napping). This followed the NASA cockpit napping study which revealed that the frequency of microsleeps was vastly reduced by the use of controlled napping. The procedure was put in place with the support of the Civil Aviation Authority of New Zealand, and allows the crew to have a fatigued crew member undertake a 45 minute period of napping, while strapped into the crew seat but wearing eyeshades and ear plugs. Before commencing, the other crew members undertake a briefing, agree on the time to wake the napping pilot, and brief the cabin crew. No course changes, altitude changes or fuel transfers are permitted during this period. Napping is permitted for 4-, 3- and 2- person crews. When the policy was introduced it was front-page news but represented a realistic approach to an undeniable problem. The procedure is employed relatively infrequently but is available as a tool to prevent serious pilot incapacitation.

Prior to introducing cockpit napping, and in order to satisfy regulatory requirements, the Flight Crew Fatigue Study Group (FCFSG) was formed and the fatigue management programme began. The underlying philosophy of the programme is that it would be data-driven rather than industrially motivated; the key to achieving this is that it is a joint initiative by management and union groups. The concept is essentially of placing unbiased data at the centre of what could otherwise be a tug-of-war over flight and duty times. The data from studies is open, in that the entire set of (de-identified) data is available to inspection by pilots, and participants can also examine their own personal data on request. A key component is external validation, and the group's work is overseen and audited by a small panel of recognised researchers from around the world with expertise in the area of fatigue and performance. This overseeing panel visits annually to critically review and guide the work of the group.

The FCFSG itself comprises a representative from each of the three active pilot union groups, a further representative of the Flight Engineers' union, two representatives of the International Airline Flight Operations management, and the technical members: the Scientific Advisor who is Associate Professor in Psychiatry and Behavioural Science at the University of Auckland; the Airline's Chief Medical Officer (the author) and the Chief Nursing Officer. The group meets monthly. One of the prime tools of the group is the fatigue report form. A system is in place for any pilot encountering excessive levels of fatigue to report this, confidentially if required, along with possible causes and remedies. The forms (about 3-4 a month on average) are passed to Flight Operations management for possible immediate action or comment, and then collated by the FCFSG to look for patterns and particularly for problem tours of duty.

A further function of the group is education. The group has had input to several levels of training in fatigue, sleep and fatigue countermeasures in the airline. This includes *ab initio* training of new pilots, recurrent training of pilots, and large open seminars to aviation industry audiences. There is also training of cabin crew and a future possibility is training of roster-writers in the logistics section of the International Airline. Some pilots have been trained with the NASA-AMES Fatigue Countermeasures Team and NASA educational material including workbooks and videos have been used in various parts of the company.

The central and most visible role of the group, however, is carrying out fatigue studies on operational tours of duty. The methodology developed allows studies to be conducted without an experimenter on board and has proved sufficiently successful that it is likely to stay. It appears to have earned the confidence of both management and the pilot population. When a particular tour of duty is identified as having a high potential for fatigue, either through generating fatigue reports or through concerns being raised by either management or union groups, it may be recommended for study by the FCFSG. Data is then collected over a period of months for analysis and recommendations to management.

Subjects are those rostered to fly the duty in question, and no selection criteria are applied other than to exclude pilots who also are involved in management, more because of their different flying hours than because of concern about impartiality. Tours which are non-standard because of some sectors being passengered, or extra crew members undergoing training, tend to be excluded. The pilots are advised in writing in advance of the study commencing, then when rostered are approached by telephone by a member of the study group. Most are known personally by the group members, and know that they are quite free to decline the invitation to participate; in practice this rarely occurs. Most of the crews studied are three person crews; if the Captain declines, the group would (unless persuaded otherwise by the Captain) not test any other crew member on that flight.

The measuring tools employed are simple. Whenever possible, participants are given a wrist actigraph (figure 1) to wear for the duration of the study. These devices are essentially accelerometers and measure wrist activity; this provides a good index of the timing and quality of sleep. Recording this from prior to departure provides information about both major sleep periods and napping, in-flight and on the ground. The earlier studies used an "Actilume" device which simultaneously measured ambient light. This provided little extra information and had the disadvantage that the device was heavy and bulky, needing two straps to fix it to the wrist. We have now moved to a "Sleepwatch" which is smaller than a normal wristwatch and is much better tolerated by the pilots.

The other tools are applied in-flight. The first is a questionnaire which employs standardised subjective ratings (figure 2). These are: Fatigue Visual Analogue Scales which allow pilots to mark how they feel on a scale between two extremes, and the position of the mark is measured; the Profile of Mood States, a standardised set of words which are scored between zero ("Not at All") and 4 ("Extremely") by the pilot; and lastly the Stanford Sleepiness Scale, a score between one and seven where each number corresponds to a word picture of the individual's fatigue feelings at the time. The three questionnaires take only a few minutes to complete.

The subjective ratings are coupled with the objective test which is the Psychomotor Vigilance Task (PVT) developed by the University of Pennsylvania at Philadelphia. This is a box the size of a novel (figure 3), which the pilot sits down with for ten minutes. At semi-random intervals a light emitting diode (LED) counter display illuminates and the pilot extinguishes it as quickly as possible by pushing a button using his/her thumb. The counter then briefly displays what the reaction was in milliseconds, before illuminating again. Over ten minutes about 100 tests are run. The device stores all of the responses for subsequent analysis. This device has been validated in a number of settings as a measure of performance and alertness. Various parameters can be measured but the most useful to date has been a score of the number of "lapses" which are defined as responses exceeding 500 milliseconds. The transformed lapse scores in our studies have correlated very well with the subjective ratings.

Having agreed to participate, crews are briefed in person by a member of the study group prior to leaving New Zealand on the tour of duty. They are also given written notes on how to carry out the testing, and there is a facility for them to contact group members by long-range communications if necessary, but this has not yet been required. Most of the previous experience with the PVT overseas is from studies with an experimenter present during testing. The first study performed by the FCFSG was on a Tour of duty Auckland-Narita-Christchurch-Auckland with an experimenter present on some sectors. The data demonstrated no difference attributable to the experimenter. Subsequently, study group members have flown with some of our crews and concluded that little value was added by having an experimenter present, provided that the pre-flight briefing was thorough and carried out using a standardised checklist.

Early in the programme, we contribute to a NASA-AMES study examining bunk-rest. This was done on a Taipei-Brisbane tour. This work has not yet been published by NASA.

In-house testing has subsequently been carried out by the FCFSG on several routes. The first major jet-lag study was on a 747-400 tour of duty Auckland-Los Angeles-Frankfurt-Los Angeles-Auckland. This study was aimed at ensuring that the measures were consistent, determining whether the methodology could be used self-administered, also incorporated measurement of salivary melatonin. The fatigue results showed good consistency between the different measures, and indicated that self-administered testing including the PVT was feasible. Collection of salivary melatonin presented many logistical challenges and results were not conclusive. However there was some evidence that there was circadian disruption before the tour began, and this increased through the tour.

The next study resulted in a tour of duty being altered. There had been frequent fatigue reports from the Taipei-Brisbane-Auckland flights on the Brisbane-Auckland leg. These "tag" flights which have a short flight (3:10 hours Brisbane-Auckland) piggy-backed on to a long one (8:50 hours Taipei-Brisbane) are often cited as being associated with high levels of fatigue. The results of the testing, in particular the PVT lapse scores, were within acceptable limits, but not by a large margin, and the trend toward high levels of fatigue was clear (figure 4). As a result of these data the company reviewed the crewing of this tour and pre-positioned slip-crews in Brisbane to operate the sector to Auckland.

The next study was of what appeared to be a very similar tour, Narita-Nadi-Auckland. The flight times are similar as are the time zone changes (figure 5); the aircraft was a B747-200 whereas it had been a B767 on Taipei-Brisbane-Auckland, but both were three-pilot crews. The results did not display serious levels of fatigue. There is very good agreement with the two studies in terms of trends, both in subjective and objective data, but the Narita journey was significantly better (figure 6). The most likely factors identified to explain this were that the hotel in Narita is quieter and crews reliably report achieving better rest there; and that the journey between hotel and airport was significantly shorter (15 minutes versus one hour). This comparison points out the potential advantage of a scientifically based approach rather than one considering only flight and duty times, which would have treated these two tours as equivalent.

A subsequent study focussed on a tour which had not generated any fatigue reports and was intended to form a baseline for comparison with future data. This was Auckland-Perth-Auckland on B767, involving an afternoon flight out and an overnight flight back across five time zones. Results showed that crews slept well in Perth, that they reliably achieved an afternoon nap there prior to the evening departure, and that performance though deteriorating through the night was maintained at acceptable levels (figure 7). Though numbers were too small for reliable analysis of sub-groups, it appeared that there was no difference between those who stayed for two nights in Perth and those who stayed five nights. In this study there were two extreme outliers in the PVT data which were excluded from the analysis, although the conclusions would be unaltered if they were included.

More recently we carried a study for one of Air New Zealand’s subsidiary airlines which operates an aircraft with a two-pilot crew between New Zealand and Australia. The sectors of interest departed New Zealand early evening and returned around dawn the next morning. As they also operate the same routes during daylight this study provided an excellent opportunity to compare daylight and night-time data from the same subject group. The results (figure 86) show graphically the difference in performance from operating during the “window of circadian low” between 0200 and 0600 – and this group of subjects does not overnight outside New Zealand so time zone shifts are not a factor. Fatigue results did not reach worrying levels but recommendations were made to improve the chances of pre-flight rest and enhance performance towards the end of the overnight flights.

The most recent study looks at a tour Auckland-Los Angeles-London-Los Angeles-Sydney. The first layover in Los Angeles is one day, and the London layover can be either 24 or 48 hours. Many pilots favour the shorter layover, preferring to sleep in daylight and be active at night; they feel that a second night in London provides poor sleep and an unwanted circadian shift. Conversely, others feel that the second day allows for better recovery before the London-Los Angeles sector and a “second chance” at getting good rest. The study is designed to examine any differences in performance on this sector depending on whether the crews had one or two nights in London, and to look at whether individual preferences relate to subsequent performance.

One of the disadvantages of the PVT is that it is reasonably bulky, it needs recharging at foreign ports, and the data analysis is quite complex and time-consuming. We have been investigating alternative technologies for some time and are in the process of validating the Pilot Alertness Test –

Air New Zealand (PATANZ). This uses a handheld electronic "Personal Digital Assistant" to apply similar types of test; the one we have used is the 3Com "PalmPilot" (figure 97). Each crew member is given one PalmPilot at the pre-flight briefing. The same subjective ratings as were on the questionnaire are programmed into the device. The objective testing uses the PATANZ which is, rather than a simple reaction time, a choice reaction time. The signal to be extinguished is on an LCD display and the pilot presses different buttons to extinguish it depending on where the signal appears. Testing has begun both in-flight and in other settings and no operational decisions will be taken on the basis of the PATANZ until it is validated.

The problems we have encountered on the programme have been logistical in most cases. Ensuring that equipment is recharged, transported and returned to us on time, then downloaded and analysed in time to be quickly despatched again can be reasonably labour-intensive, not to mention the requirement for in-person briefings of crews at the airport at odd hours. Many problems have centred around issues such as batteries, power plug adaptors, and customs clearances – but these are the stock and trade of gaining experience. There is a need to ensure that briefings are thorough and specific to avoid misunderstandings. There are also some challenges with interpreting borderline data, and criteria for excluding the data of extreme outliers from the analysis. In these areas having external advisors overseeing and auditing the work is invaluable.

The long-term aim of the programme is to achieve flight and duty time limitations that are considered safe and acceptable on the basis of reliable data. It is hoped that in future, the group will be involved pro-actively in advising on tours of duty before they are introduced, rather than testing only after fatigue reports are received. The PATANZ test, when validated, should streamline testing and hasten the production of reports and recommendations. The programme may well extend to involve cabin crew and maintenance personnel, in whom fatigue has an important bearing on safety.

The intent of this paper has been to outline a programme that is already in operational uses as a tool to manage the thorny issue of flight-crew fatigue in a competitive market-driven environment. The strength of the programme is that it is founded on data rather than interests, and enjoys support from both airline management and pilot union groups. Combining fatigue reporting and education with both subjective and objective testing is proving its worth in an operational setting. A decision was made this year that rather than attempt to market this system, we would open it up to the public domain for the betterment of flight safety in the International Aviation Community.

References: