Fatigue at the Top of the Drop

Review of a Fatigue Risk Management System in a Commercial Airline Setting

David Powell
Chief Medical Officer, Air New Zealand Limited

Introduction

This is an example of a Fatigue Risk Management System in practice in a commercial airline. The presentation will outline the development of the programme and focus on some recent studies performed, looking at the benefit of an extra night layover, and the benefit of an extra pilot. It will then describe a "Top of Descent Survey" recently completed, and look to future developments in the fatigue and alertness work.

Few would dispute that optimising the alertness of aircrew is a vital component of flight safety in commercial aviation. In recent years there is growing interest in replacing more prescriptive systems of flight and duty time limitation, with more scientifically based approaches, more latterly called Fatigue Risk Management Systems. We have been operating and refining such a system for several years as part of our compliance with the applicable Civil Aviation Rules on alertness management.

Background

By virtue of New Zealand's location on the globe, it is inevitable that Air New Zealand is a long-haul airline. It operates about 40 jets (8 Boeing 747-400, 10 Boeing 767, 7 Airbus 320, and 13 Boeing 737-300) as far as UK, crossing 12 time zones. Air New Zealand established a collaborative group to monitor aircrew fatigue some eight years ago. Since then the group has conducted over 20 studies on operational tours, and has probably the largest set of data on operational pilots. The team, called the Crew Alertness Study Group (CASG), has representation from aircrew Union groups, management, scientific and medical, as well as rostering & scheduling representatives. It is this joint membership and representation that is the strength of the system. The Group's primary role is to gather and analyse scientific data on fatigue, to assist management with making decisions based on science. In this way it is a departure from the more traditional "tug-of-war" approach between airline management and unions.

The scope of the group was expanded more recently to encompass cabin crew as well as pilots. As well as data collection and analysis, the group also provides advice to management and education to aircrew on fatigue. One of the key strengths of the group has been the balanced membership consisting of aircrew union representatives, management, and scientific/medical members. At all times it is stressed that the function of the group is to provide data, not to make decisions. Another key feature is that from the group's inception, there has been an external review body, the International Alertness Advisory Panel (IAAP) whose role is to review and oversee the work of the group, in particular the interpretation and conclusions of the studies, to ensure scientific rigour and validity. Figure 1 shows the structure of the system, and the current IAAP members.

Fatigue Reports

Indications of trips of concern are gained from aircrew fatigue reports; pilots and cabin crew are encouraged to submit reports (Figure 2) when encountering undue or excessive fatigue while flying – and also to indicate causative factors and suggested remedies. The reports, after they have been received and actioned by management, are passed to the CASG which monitors trends. If requested by the reporter, a report will be de-identified, but in most cases this is not asked for.

Operational Studies

If a particular trip (Tour of Duty, TOD) is identified as a concern for alertness, a study may be initiated. This may be on the basis of fatigue reports received, or requested by management or by the group itself. Typically some 20-30 pilots flying the TOD, will be asked to participate in the study over a period of several weeks. Generally almost all pilots agree to be tested. A briefing is conducted prior to departure and equipment issued and demonstrated; testing is done on board by the crew members themselves, without an experimenter on board.

Tools

While our earlier studies were done using PVT (Psychomotor Vigilance Task) developed by Dinges et al (figure 3), we subsequently developed and validated our own tool based on a PDA (personal digital assistant), shown in figure 4. After entering some basic personal data, the pilot is then asked to complete testing at regular intervals, typically two-hourly, during the cruise phase of the flight. Each test consists of a series of subjective ratings (Visual analogue scales of fatigue, Karolinska Sleepiness Scale, Samn-Perelli Alertness Rating) followed by a 10-minute alertness test. In this test, a shape is presented on the screen of the PDA at pseudo-random intervals, some 100 times; the pilot's task is to press the corresponding button on the PDA as quickly as possible, and the response times (in milliseconds) are displayed and recorded (as well as any errors). The measure we have found most useful is the simple mean reaction time in milliseconds. Figure 5 shows some of the screens from the PDA.

In addition, the majority of studies also involve actigraphy. A wrist actigraph (figure 6) is worn by each crew member throughout the tour of duty and records the amount of motion as a function over time; the output from this provides a reasonable indication of when sleep is occurring.

Validation

This task was validated in three stages. In the first, we sleep-deprived 12 volunteers for 31 hours, free of alcohol or caffeine, and serially tested them with both our alertness test and the PVT, using a counterbalanced design. Comparative results are shown at figure 7. The trends in mean reaction time were as expected, mirroring the trends in the transformed lapse scores on PVT.

The second validation stage was a field trial in which we tested crews flying Auckland-London-Sydney, testing one group with PVT and the other with our PDA-based test. Similar results were obtained (figure 8). Again, the trends observed, and conclusions reached, matched those associated with the alternative task.

Finally, the test was used in a study with QinetiQ on Brittania Airways flying Haj pilgrims from Indonesia to Saudi Arabia. Here, large numbers of pilots were tested in-flight on the same route leaving at different times of day; thus the only factor varying was the circadian phase. Here again, the observed variation in reaction time varied as expected. Figure 9 shows some indicative data).

Results

Figure 10 shows a summary of all studies completed to date; it will be readily seen that in about a third of cases, a change in the scheduling or crewing has followed on from the data (and recommendations) being provided to management.

In all studies, subjective fatigue and mean reaction time increase steadily through each sector. The reading of most interest, naturally, is the final test before top of descent, when fatigue is both at its greatest level, and of greatest importance. Most studies have focused on the level of alertness on this final test.

An example – Los Angeles Studies

During 2002-3 we completed a series of tests on Boeing 747-400 routes ending with a Los Angeles–Auckland (LAX-AKL) overnight sector. This sector is 12 hours 45 minutes and previously was usually crewed with three pilots – a Captain, First Officer and Second Officer. The Second Officer is a "cruise pilot" not in an operational seat for departures or approaches, but employed to allow the Captain and First Officer to take in-flight rest (each pilot taking turns using the flight deck bunk). It was our longest three-pilot sector, is usually flown in darkness the entire journey, and was a frequent source of fatigue reports. On the rosters it was preceded by two different flights – one of them Sydney-Los Angeles (SYD-LAX), a four pilot sector largely departing Sydney around midday and arriving in the morning, and the other Auckland-Los Angeles (AKL-LAX), another three pilot journey departing Auckland at night but arriving in the early afternoon (local time). Times are shown in figure 11. In both cases the pilots arrived in Los Angeles the day before they departed (ie one night of layover), but in the case of SYD-LAX they arrived in the morning. The concerns about fatigue were emanating disproportionately from the SYD-LAX-AKL trip. Testing indicated that the level of alertness was borderline (figure 12).

Los Angeles time was 19 hours behind Auckland (in other words, 5 hours ahead but the previous day). Actiwatch analysis indicated that most pilots were taking a long (around 4 hours) sleep on arrival, as would be expected given that the arrival corresponded to around 4 am New Zealand time. At local night time, the circadian rhythm discourages sleep until about 0300 local time; but hotel noise tended to wake the pilots at about 0800 or 0900 local. The next naturally sleepy time would be about 1800 local, but this is when the pilots would be getting ready to leave for the return flight. The entire journey thus afforded little opportunity for effective layover rest, and the level of fatigue observed on the homeward journey is perhaps not surprising.

Effect of an Extra Pilot

Subsequently changes were made to the rostering of Los Angeles duties. Further testing was done when some of the trips were crewed with an extra pilot. The fourth pilot, another second officer, allows two pilots (one second officer and one other) to be on rest at any given time in the cruise phase, and hence the in-flight rest opportunity is greatly increased. The results (figure 13) showed a marked benefit from having an extra pilot.

Effect of an Extra Night

Also around this time, a decision was reached to try scheduling an AKL-LAX-AKL trip. This was again three pilots with one night layover in Los Angeles. Testing of this (figure 14) showed the alertness to be lower than on the SYD-LAX-AKL flights. A decision was made to introduce an extra night of layover, and to do further testing following this change. Somewhat surprisingly, although the pilots reported a great improvement, measured alertness was little better, if at all, with the extra night in Los Angeles (figure 15).

In summary, while the two trips are not identical, for a problematic duty we were able to demonstrate the relative benefit of an extra night layover (marginal) compared with an extra pilot to facilitate in-flight rest (significant benefit). This has had implications for planning of new routes developed since then, including Auckland-San Fransisco and Christchurch-Los Angeles. Note that in all cases, the other subjective ratings (Karolinska Sleepiness Scale and Samn-Perelli) showed almost identical results to the Visual Analogue Scales.

"Top of Descent Survey"

Some concerns were expressed to us that because of a reluctance to report fatigue, it might be misleading for us to base our assessment of tours of duty on the numbers of fatigue reports being received. In response to this concern we undertook a fleet-wide "top of descent survey" over a three month period. Each survey form took only a few seconds, and asked the pilots to rate their alertness using the Samn-Perelli scale and an "Alert-Drowsy" Visual Analogue Scale (figure 16). Some 9000 alertness ratings were received and analysed. This allowed us to rank the different trip patterns according to alertness, and gives a comprehensive comparison database for future trips of concern. Figure 17 shows some of the results.

Ultra-Long-Haul (ULR)

Although not planning ULR operations in the medium term, Air New Zealand has participated actively in the Flight Safety Foundation/Boeing/Airbus workshops on crew alertness for ULR operations. Four of our personnel contributed including two on the steering committee. We believe that a scientifically based approach is vital to making such decisions as how many crew members are necessary. The PDA-based alertness test is being used currently on testing by QinetiQ on Singapore Airlines' new ULR operation with the A340-500 on Singapore-Newark-Singapore.

Developments

Recently we examined the feasibility of shortening the alertness test on the PDA to make it less of an imposition for aircrew being tested; our analysis indicates that 50 choices (with a duration of around 5 minutes) will discriminate the performance level almost as well as the full 100-choice test. We hope to make this shortened alertness test available for free download on the internet in the near future.

The future:

All of the work described thus far has focused on trips already in existence. Our goal is to be able to determine the trips which are likely to be problematic in advance, before they come in to existence. To this end, we have worked for some years with QinetiQ on developing their SAFE model for use as part of our trip and roster design. Figure 18 shows a typical output from the model. Data from a number of our studies have been used by QinetiQ to validate and refine the model, particularly on journeys crossing large numbers of time zones. In turn, QinetiQ and Air New Zealand, under a Teaming Agreement with the UKCAA, have developed a capability the model to be integrated directly into the optimiser which creates of new rosters and schedules, checking for fatigue potential before the rosters come into being.

Making it Real

One of the limitations of most fatigue research in aviation is the lack of a tangible link with meaningful operational performance measures. Although there are reports of accidents linked to fatigue, this is too extreme a measure and there is a need for lower-level decrements in performance to be linked to alertness. On the other hand, while tracking error on simulator tasks can be demonstrated, the link between this and safety outcome is tenuous.

One exception is work by Russo who looked at performance on an in-flight refuelling task on USAF bomber simulators. Alertness was measured with PVT, while there were measurements of accuracy, errors and efficiency at detecting peripheral visual signals. A clear relationship was demonstrated between ability to conduct the task and detect signals, time awake, and PVT lapse scores. Lapses, errors, and visual neglect all increased steadily beyond 18-19 hours continuously awake.

The next area which we would like to explore is the analysis of performance as demonstrated on the outputs from the Quick Access Recorder within a Flight Operations Quality Analysis (FOQA) system. It is only by comparing operational performance with alertness, that we can approach an answer to the ever-elusive question – how fatigued is too fatigued?

Summary

Over several years we have provided alertness data to assist management decisions on rostering, scheduling and crewing of both pilots and flight attendants. The data has consisted of fatigue reports, in-flight studies, survey data and the outputs of the QinetiQ SAFE model. We suggest that these are the key elements of a practical Fatigue Risk Management System for a commercial airline.

References

On request from author.

Acknowledgements

Dr Keith Petrie, University of Auckland
Mr Mick Spencer, QinetiQ UK
Dr Curt Graeber, Boeing Commercial Airplanes
Captain David Morgan, Air New Zealand Limited
Captain Bryan Wyness, Air New Zealand Flight Operations (Retired)
All the members of the Air New Zealand Crew Alertness Study Group

Figures

Figure 1 – Summary of Air New Zealand Crew Alertness Programme

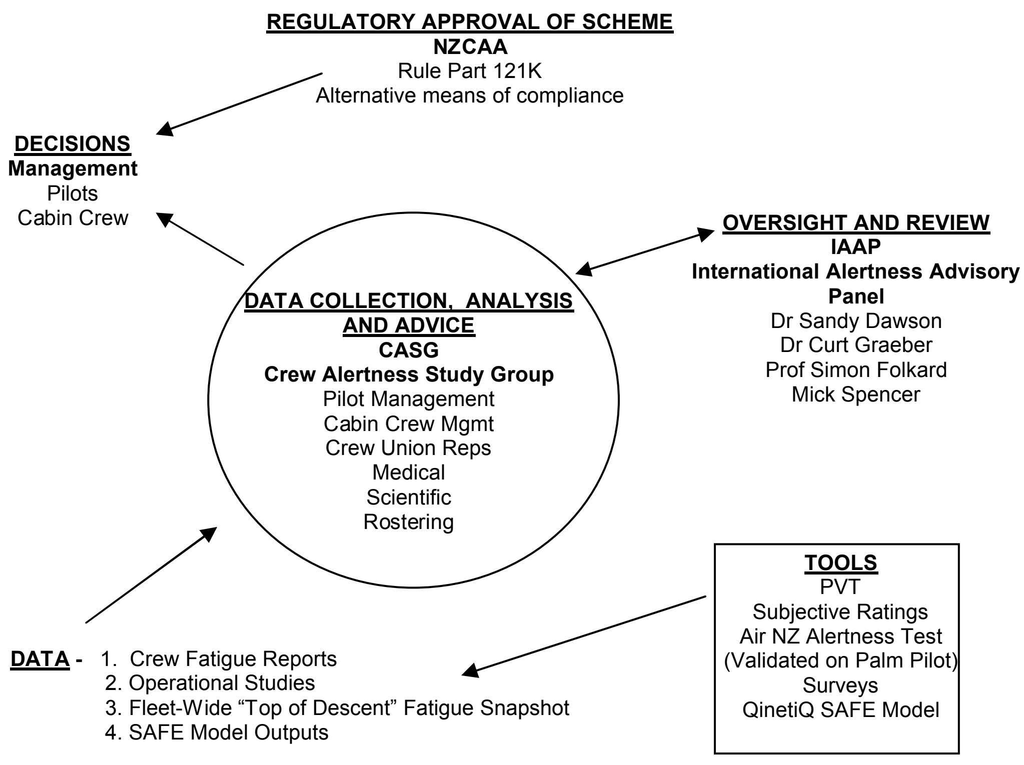


Figure 2 – Standard Fatigue Report Form

AIR NEW ZEALAND

A STAR ALLIANCE MEMBER

Fatigue Report Form

If Confidentiality required tick here ○

Name		Employee No.		Pilot / CCM	(Circle)

When did it happen?

Local Report Date		Local Report Time	
Duty Description (e.g. "LAX1287" or "AKL-CHC-ZQN-AKL"			
Sector on which fatigue occurred: FROM		TO	
Hours from report time to when fatigue occurred		Disrupt?	Yes / No
Aircraft Type		Number of Crew	Pilot / CCM (Circle) No.:

What happened?

Describe how you felt (or what you observed)

Please circle how you felt

1.	Fully alert, wide awake	5.	Moderately tired, let down
2.	Very lively, somewhat responsive, but not at peak	6.	Extremely tired, very difficult to concentrate
3.	OK, somewhat fresh	7.	Completely exhausted
4.	A little tired, less than fresh		

Please mark the line below with an "X" at the point that indicates how you felt:

ALERT ———————————————— DROWSY

Why did it happen?

Fatigued prior to Duty	Yes / No
Hotel	Yes / No
Home	Yes / No
Duty Itself	Yes / No
In-Flight Rest	Yes / No
Disrupt	Yes / No
Personal	Yes / No
Other / Comments	

What did you do?

Actions taken to manage or reduce fatigue (e.g. cockpit nap)

What could be done?

Suggested Corrective Actions

Figure 3 – Psychomotor Vigilance Task (PVT)

Figure 4 - Personal Digital Assistant (PDA)

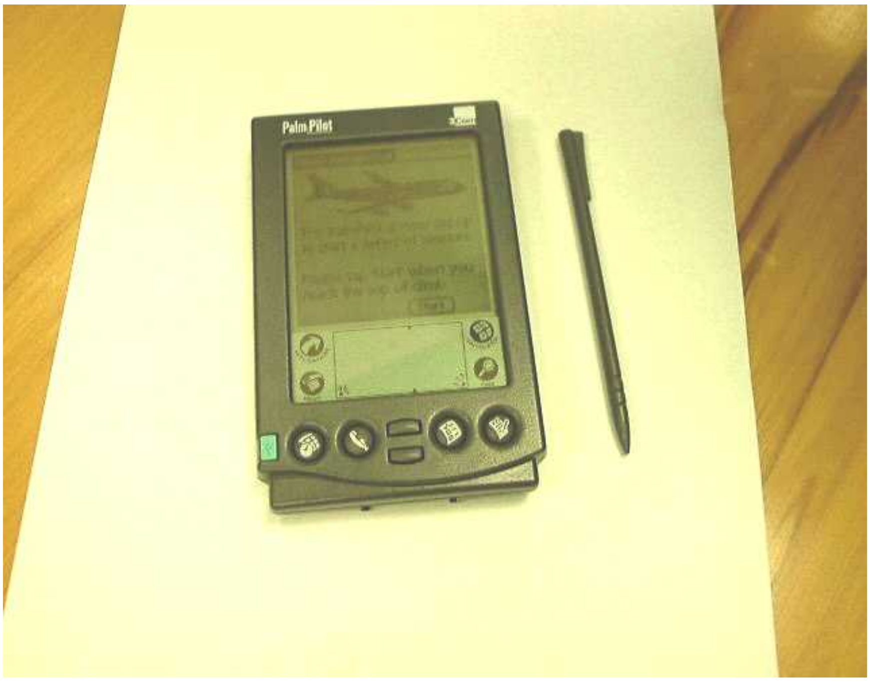


Figure 5 - PDA Screens

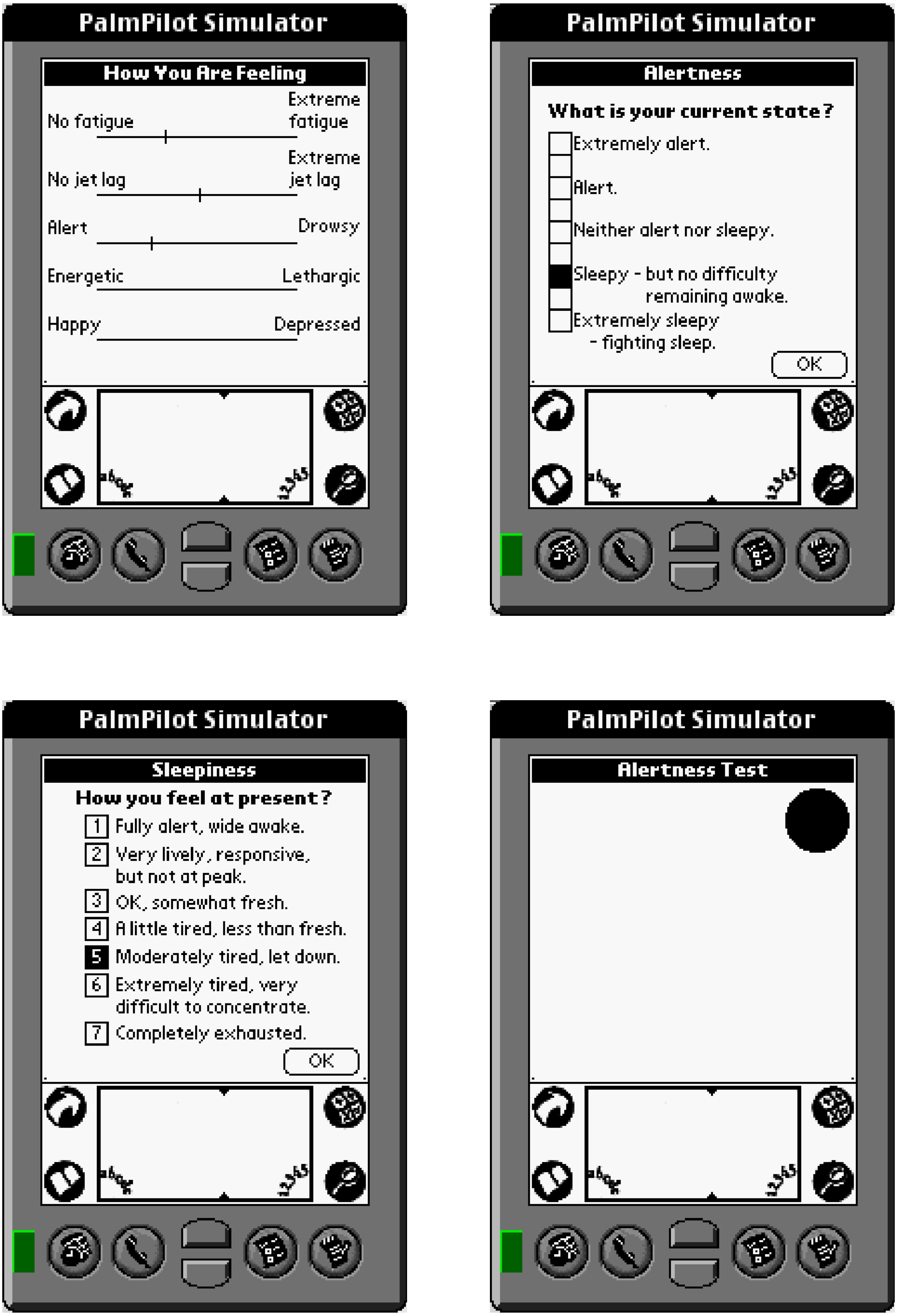


Figure 6 – Wrist Actigraph

Figure 7 – Mean Reaction Time vs PVT Transformed Lapses

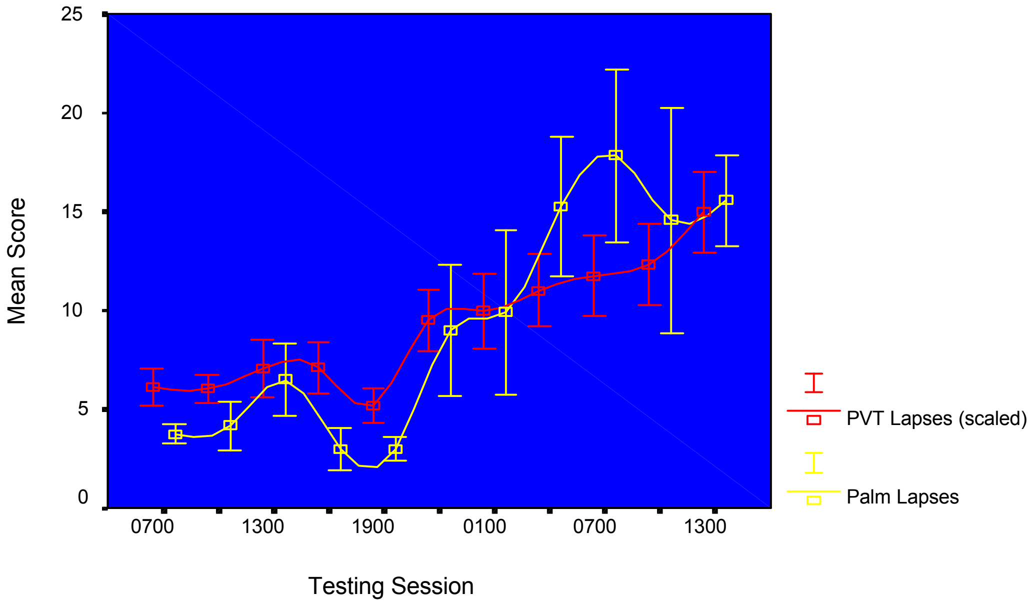
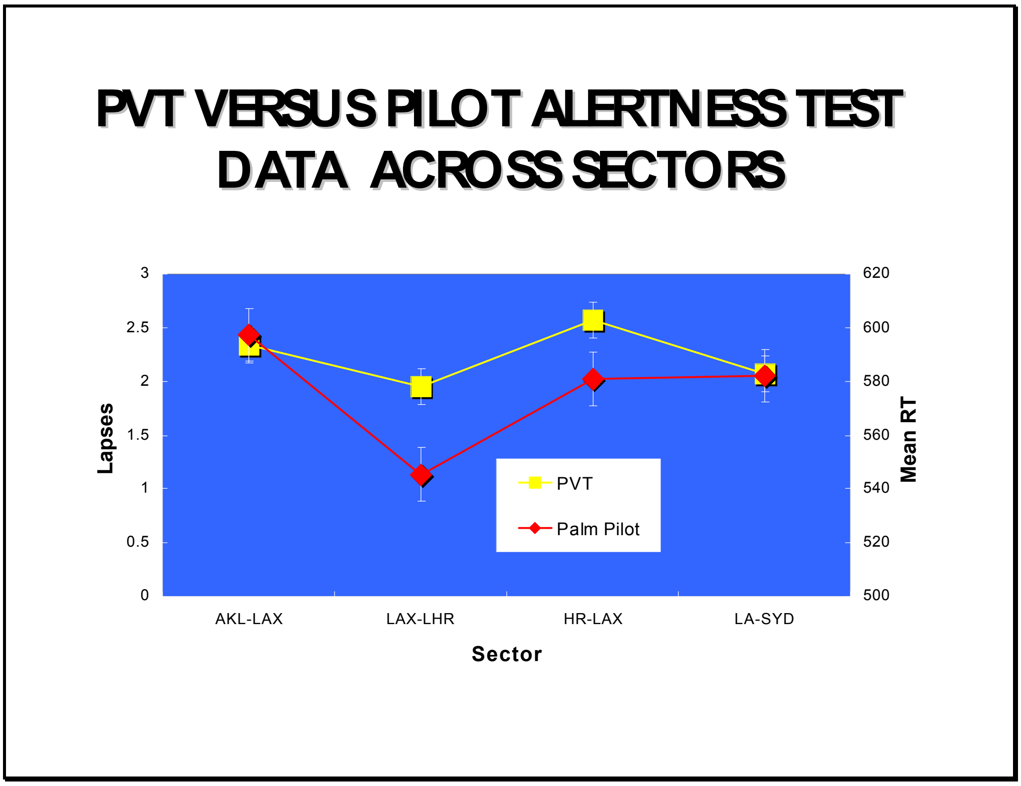


Figure 8 – Two analyses of AKL-LAX-LHR-LAX-SYD



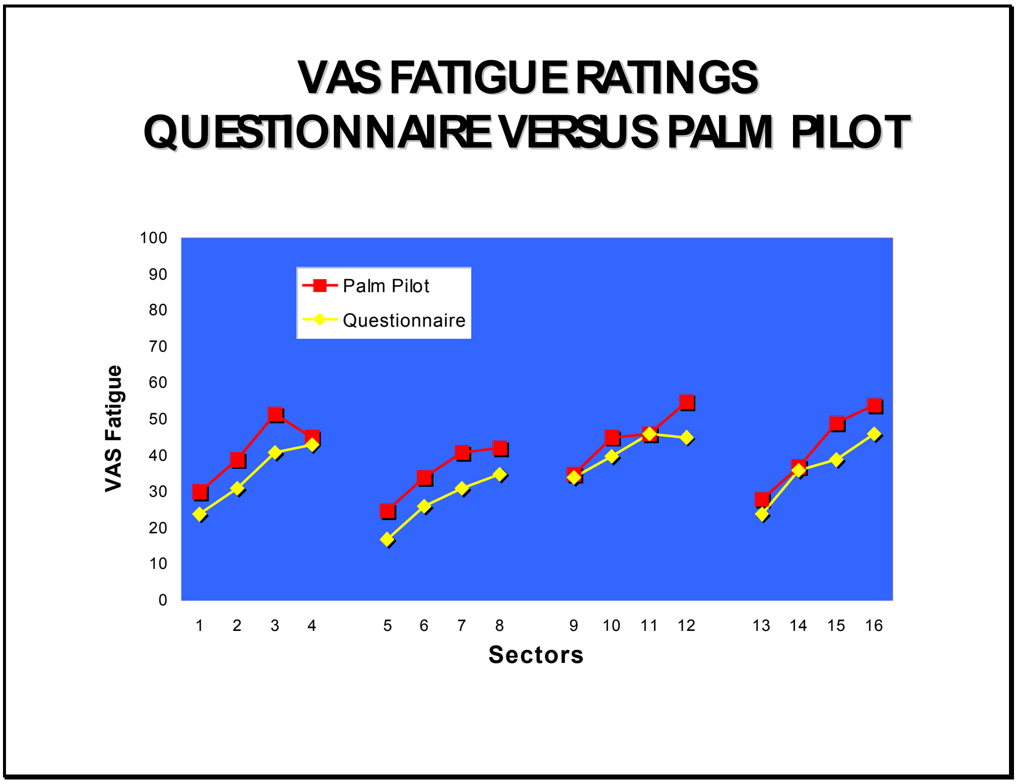


Figure 9 – Pilot Alertness Test – Haj flights, 2 hourly around the clock
Same task repeated over 980 test sessions (216 pilots)

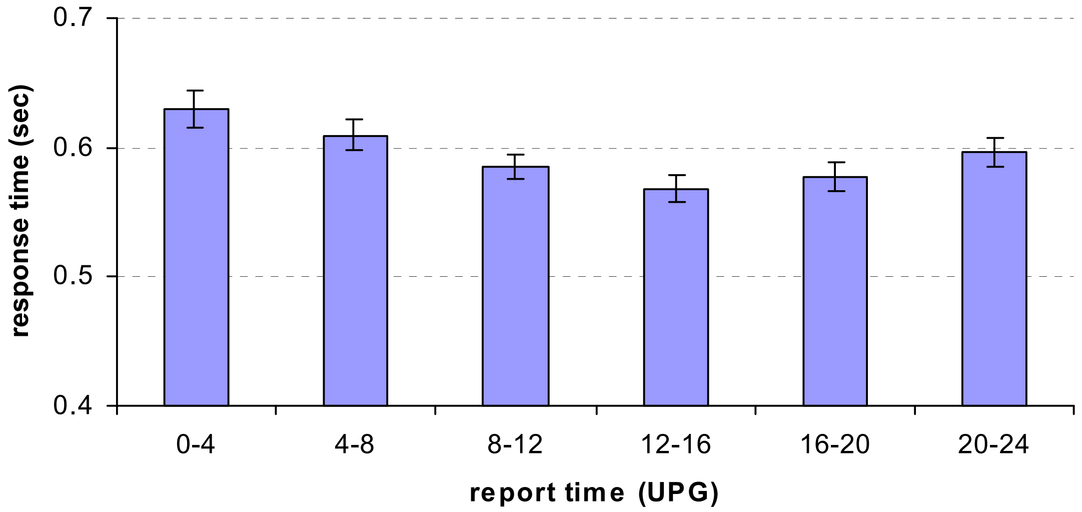


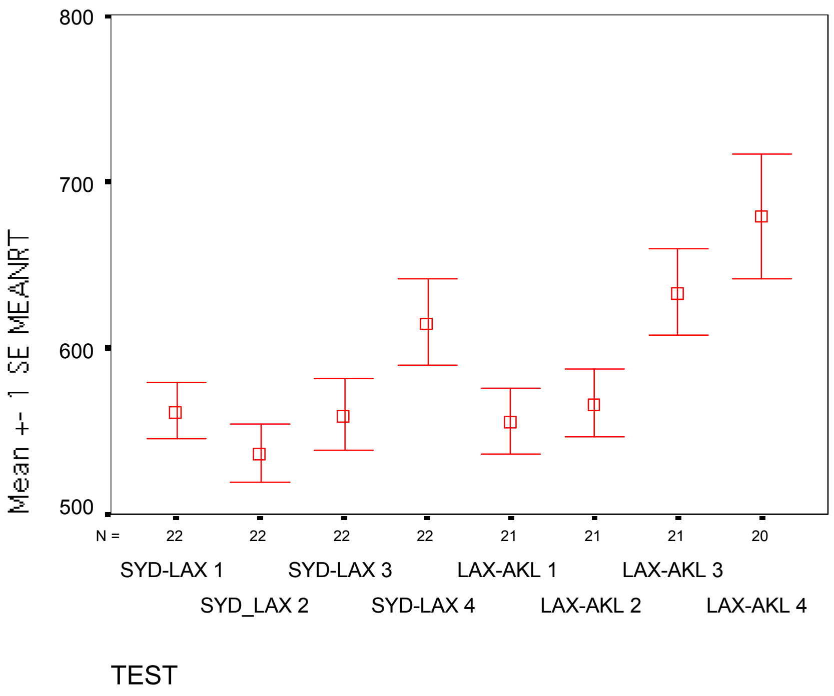
Figure 10 – Studies Completed to Date

Pilots	Cabin Crew
TPE-BNE-AKL*	AKL-NAN-RAR-PPT-RAR-NAN-AKL*
AKL-LAX-FRA-LAX-FRA-LAX-AKL	AKL-KIX-CHC-AKL
NRT-NAN-AKL	AKL-PER-AKL
NZ-Australia-NZ (Freedom Air)	AKL-TBU-HNL-AKL*
AKL-SIN-CHC*	AKL-LAX-APW-AKL
AKL-LAX-LHR-LAX-SYD	AKL-LAX-AKL
SYD-KIX-BNE-SYD (Ansett)	CHC-BNE-CHC
SYD-LAX-AKL*	
AKL-LAX-AKL*	
AKL-LAX-LHR-LAX-AKL	
CHC-BNE-CHC*	(* = Changes Made)

Figure 11 – Flight Times SYD-LAX-AKL

Departure (Local)	Arrival (Local)	Flight Time	Layover
SYD 1135	LAX 0905	13:30	37 hours
LAX 2130	AKL 0515	12:45	

Figure 12 – SYD-LAX-AKL Mean Reaction Times and Fatigue Visual Analogue Scales



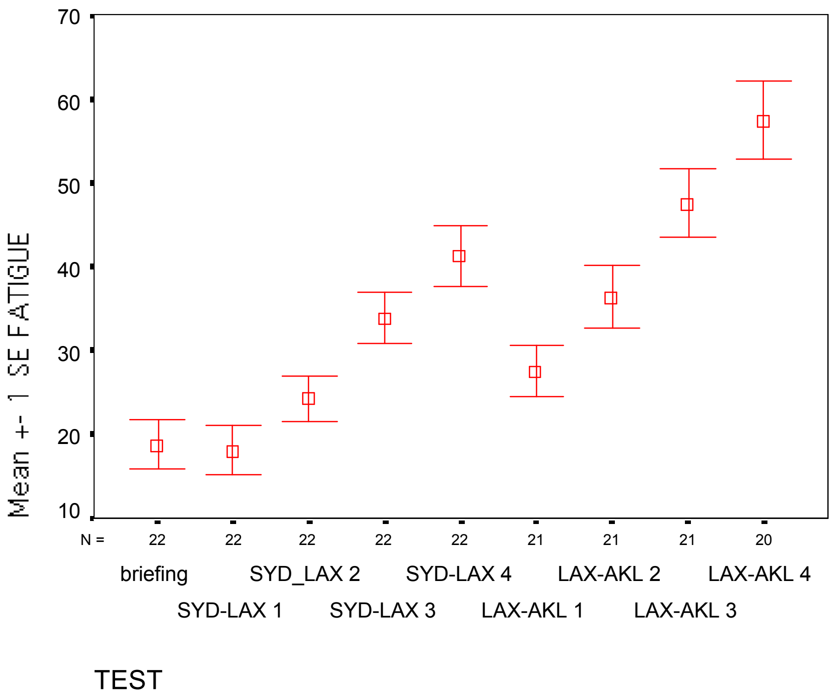
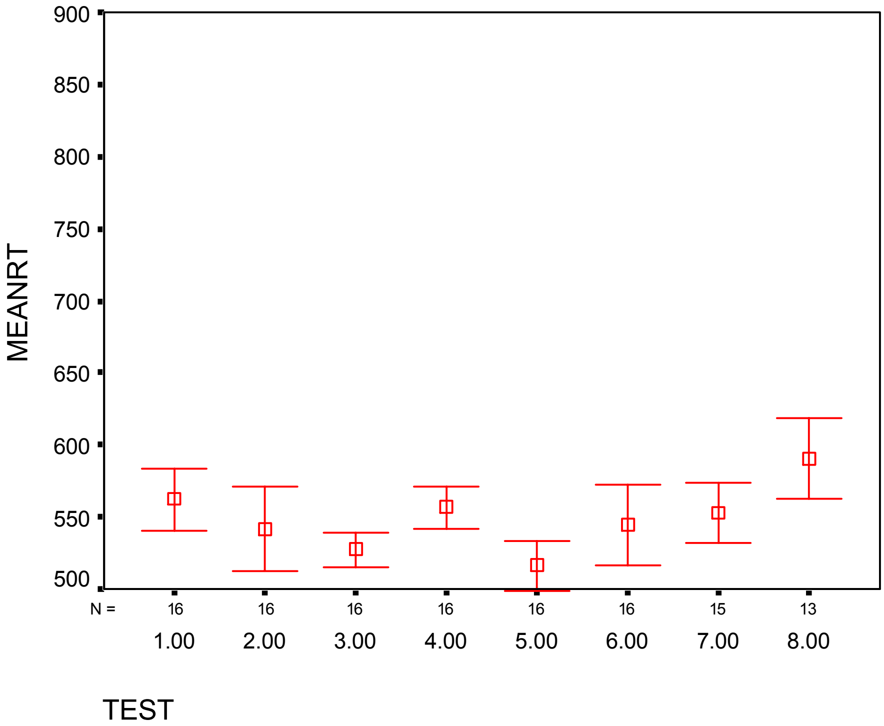


Figure 13 – Effect of a Fourth Pilot on SYD-LAX-AKL
Mean Reaction Time and Visual Analogue Scale Fatigue



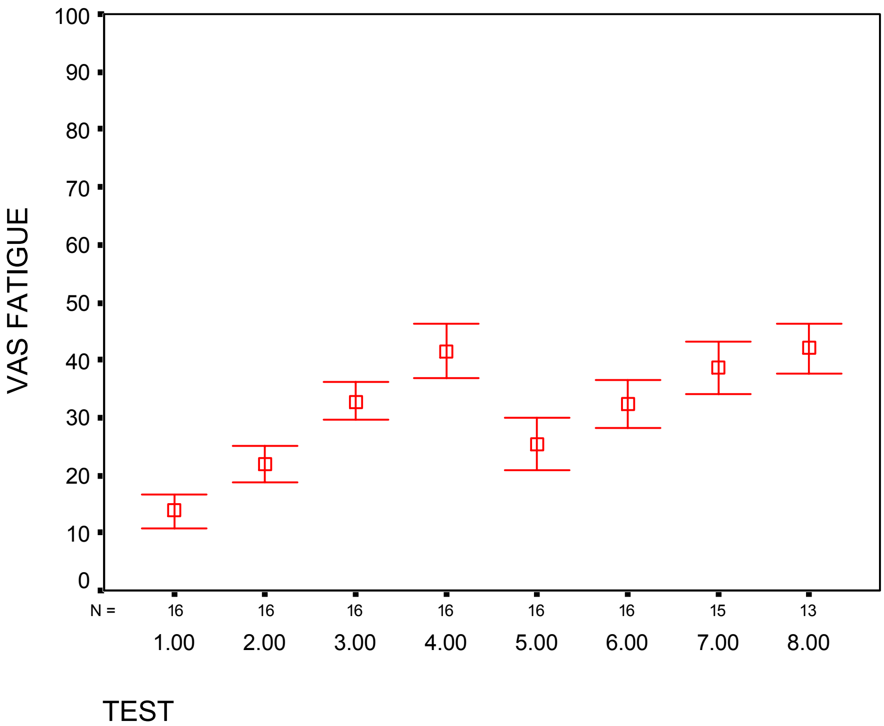
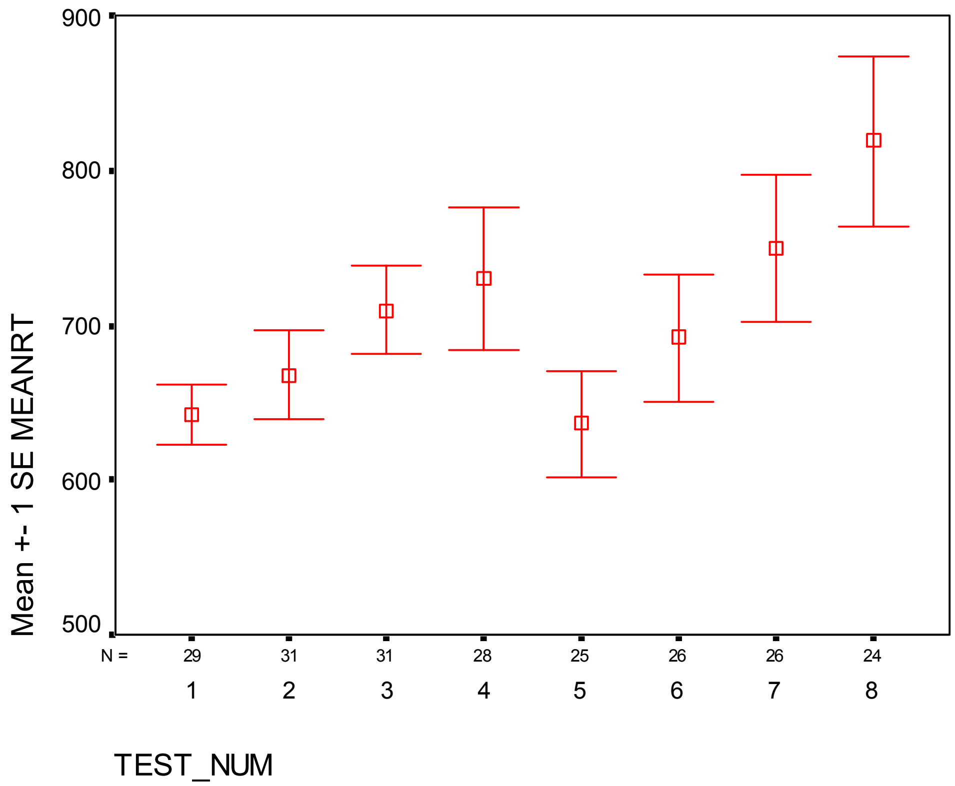


Figure 14 – AKL-LAX-AKL 3-pilot 1-night Layover
Mean Reaction Time and Visual Analogue Scales Fatigue



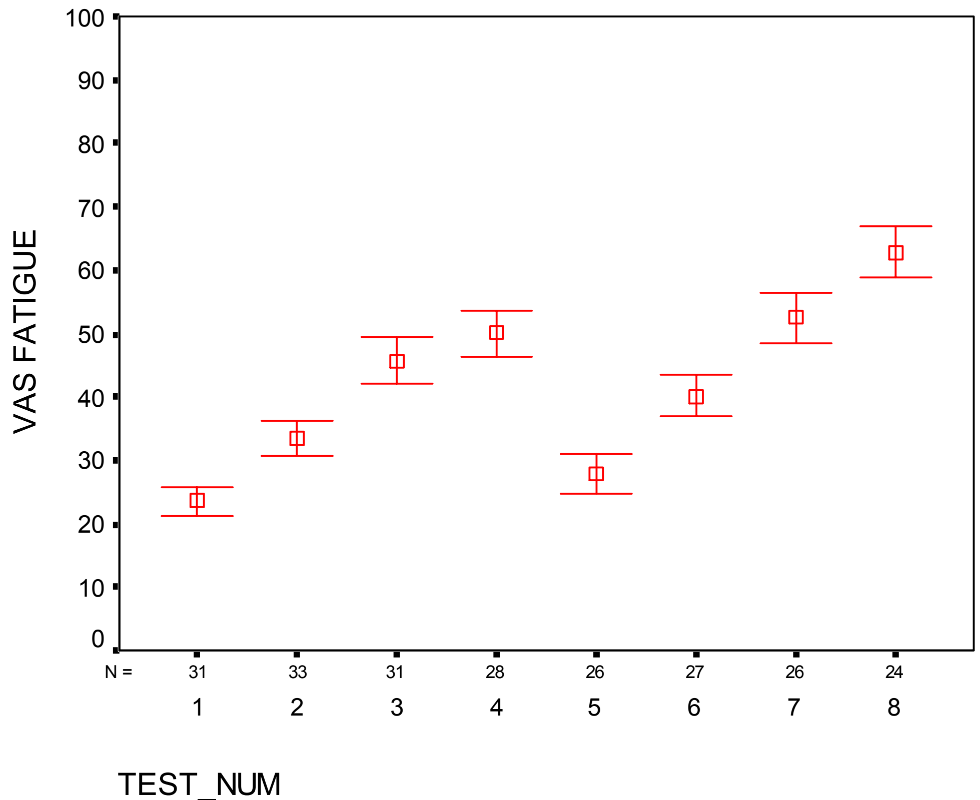
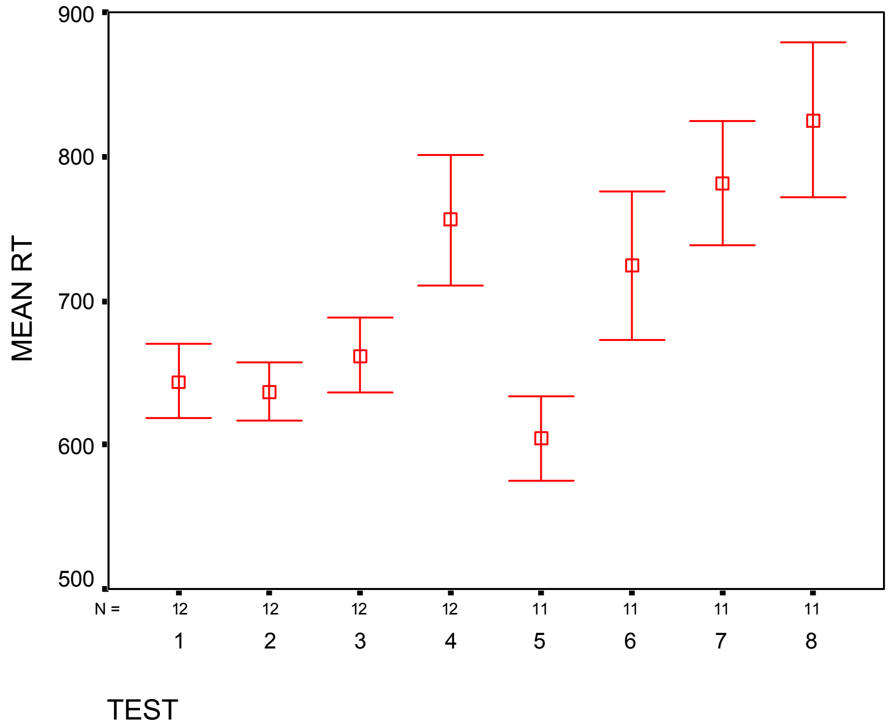


Figure 15 – AKL-LAX-AKL 3 pilot 2-night Layover
Mean Reaction Time and Visual Analogue Scale Fatigue



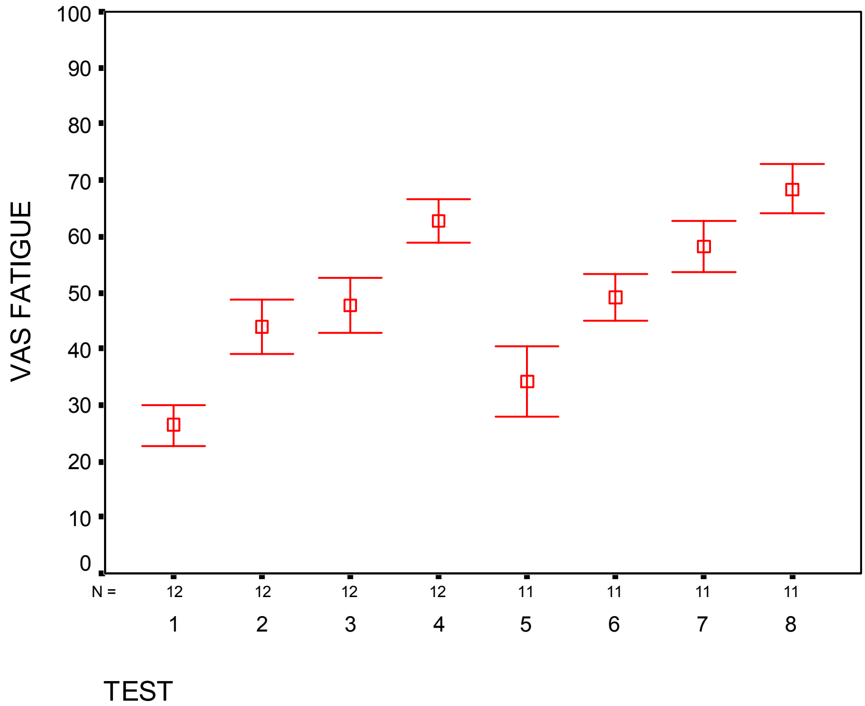


Figure 16 – “Top of Descent” Survey

AIR NEW ZEALAND

PILOT ALERTNESS REPORT FORM

Forms to be completed immediately prior to Top of Descent on last leg of duty period.

[ ] Report Time (UTC)

[ ] Time (UTC) at Top of Descent

Name the Sectors operated this duty period.

Please circle “How you feel” at Top of Descent

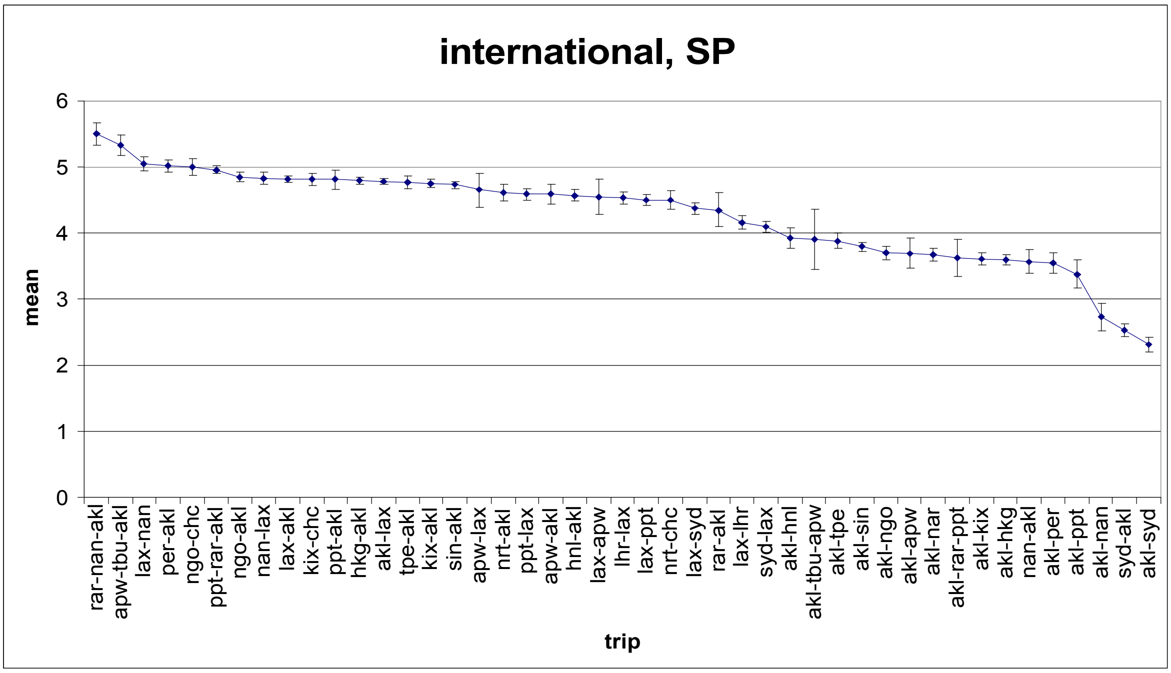
1. Fully alert, wide awake
2. Very lively, responsive, but not at peak
3. OK, somewhat fresh
4. A little tired, less than fresh
5. Moderately tired, let down
6. Extremely tired, very difficult to concentrate
7. Completed exhausted

Please mark on the line below

Alert ________________________________ Drowsy

Please place in brown envelope

Figure 17 – Results of "Top of Descent" Survey – Mean Samn-Perelli Rating and Fatigue Visual Analogue Scales



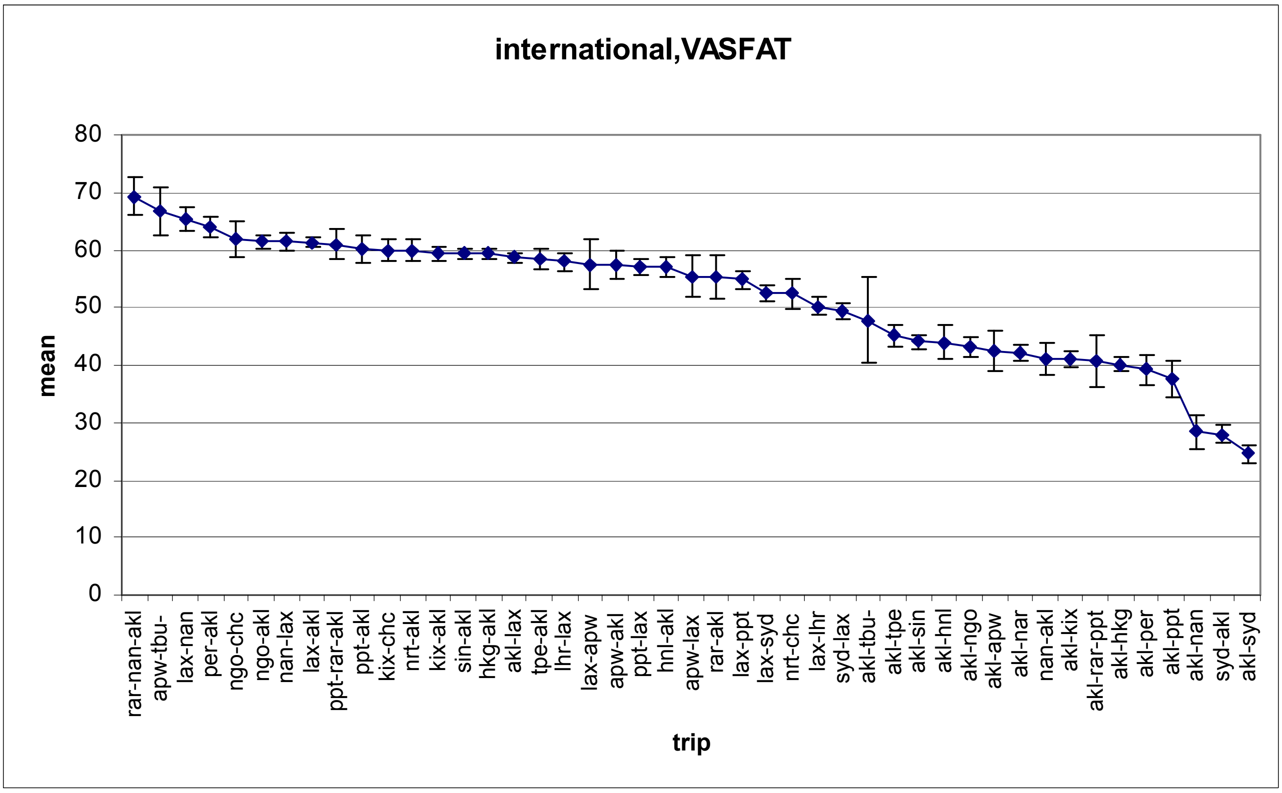


Figure 18 – Output from QinetiQ SAFE Model