But how do we measure it?

By Dr Arnab Majumdar

The evidence from various industries indicates that fatigue, primarily due to sleep deprivation, creates alertness deficits that in turn affect performance and thereby increase the chances of a serious accident. The solution therefore seems simple: eliminate fatigue by providing the necessary sleep and this should increase alertness and performance leading to the disappearance of fatigue-related accidents. End of….

Oh, if only life was that simple! Let us start at the beginning – what is meant by fatigue? In the "solid" world of materials engineering, definitions relating to the breakdown of a material due to repeated stresses form the basis of precise measurements. As soon as we get into the world of that most tricky of engineering materials, the "human being", then fuzzy definitions abound. Is fatigue something we intuitively feel or is it something that can be defined precisely, rather like the extent of progressive and localised damage to a material under a specified cyclic loading? Indeed, can it be both a subjective "feeling" and an objective "measure", and are these two related?

If definitions are unclear, how about measuring fatigue reliably? If it is about feelings, then surely all we need to do is ask individuals if they are tired and to what extent. The evidence though shows that humans are not necessarily good at judging when they are fatigued. And as if that was not enough, in the complex modern world of aviation operations, organisations cannot hope to provide an appropriate level of service based simply upon their critical personnel's feelings. However, to ignore the subjective nature of fatigue leaves the organisation vulnerable – after all, you don't want someone who feels exhausted to be sitting at controls.

And if we are measuring something objective, what should that something be? The number of hours of sleep every night, or its quality, however that is defined? Or should we look at physiological measures related to the human body? Which will be the best indicator?

The evidence though shows that humans are not necessarily good at judging when they are fatigued.

Suppose though that we can measure fatigue reliably, and acknowledging that fatigue affects performance, how can we assess if it affects safety in particular? After all, granted that fatigue may reduce performance, but will it actually compromise safety to the extent of causing incidents and accidents? Or is there an acceptable level of performance decrement due to fatigue that organisations and the public can tolerate?

These were the research questions investigated by the Air Traffic Management (ATM) Research Group at Imperial College London in conjunction with easyJet. As Lydia Hambour has highlighted in detail elsewhere in this edition, easyJet have implemented a fatigue risk management system (FRMS) which uses a variety of methods, including field studies, surveys and employee reporting in order to assess fatigue risk associated with the operating environment. This article elaborates on one aspect of the FRMS studies relating to a 3-week long trial to measure variety of fatigue, alertness and performance measures for a group of the airline's pilots.

Let us start though with a practical measurement problem. Much of the scientific literature on fatigue has involved experiments in controlled laboratory settings, as far removed from the "real world" of low-cost carrier (LCC) operations as is possible. Typically pilots can fly up to 6 sectors (i.e. routes) in a day's duty, and as an LCC, easyJet has 20-minute turnarounds between flights that are task-driven. Just in case there have also been delays during the day's flight schedule, the flight crew are under even more pressure during this turnaround. Therefore to conduct any tests to measure fatigue and its impacts must not interfere with any operational duties. On top of this, most of the flight sectors for the airline are relatively short (between 30 minutes to 4 hours), with little chance of administering tests during duties or solving any equipment glitches. Both hardware and human resource requirements thus need to be handled with care to ensure maximum cooperation from the flight crews. In these circumstances therefore, the need is to determine what tests can be feasibly conducted to provide meaningful, robust results without causing operational risks and increased costs.

An additional consideration involved setting the baseline for the subsequent measurements. Again, the controlled experiments literature provides few guidelines for the operational setting. Instead, the flight schedule for the 22 pilots taking part in the trial was designed so that they could be monitored for a 23-day period including three days off at the start and end of the monitoring period in order to establish baseline measures and to permit recovery from prior duty sequences operated. The pilots operated a specially designed schedule that involved three consecutive early start duties followed by two late finish duties (Block 1). This enabled data on performance and sleep to be collated for the first of the transition changes. Following Block 1, three days off were provided. The second duty block (Block 2) contained one further transition change and the duty sequence closely reflected the timings (Block 1). Two days off were provided following Block 2 before the pilots completed a further block (Block 3) to enable com-

Dr Arnab Majumbar

is the Director of The Lloyd's Register Educational Trust Transport Risk Management Centre (LRET TRMC), at the Centre for Transport Studies, Imperial College London and a Lecturer on Transport Risk Management at Imperial. The LRET TRMC is a centre dedicated to the analysis of risk in all modes of transport for the public benefit. Arnab has a PhD in airspace capacity assessment from Imperial College London and over 20 years of experience in air traffic management research in close collaboration with the major players in the domain.

But how do we measure it? (cont'd)

parison in performance levels following two days off with that following three days off.

And so what to measure? Kilner and Cebola have noted in this edition that fatigue is assessed by a combination of subjective and objective measures. Various studies show that the schedule affects fatigue, hence the need to monitor the schedule design, e.g. duty time duration, the number of early and late duties and number of rest days. Furthermore, both the duration (cumulatively collected) and quality of sleep was measured, as was the subjective alertness of the pilots. To measure performance, the psychomotor vigilance task (PVT) was used. This is a simple task where the subject presses a button as soon as the light appears. The light will turn on randomly every few seconds for 5–10 minutes. Both the pilots' reaction time (RT), together with an assessment of how many times the button is not pressed when the light is on, provide a numerical measure of sleepiness by assessing lapses in the subject's sustained attention. To collect this plethora of measures, specially designed workbooks were completed by the pilots, which included their subjective fatigue, sleep and alertness ratings and in addition they wore actiwatches to record their sleep and conducted the PVT tests using hand-held personal digital assistants (PDA).

Prior to any analysis, the first task involved assessing whether the pilots were uniform as a group or whether they acted as individuals. By far, uniformity was observed, thereby allowing the results of statistical analyses on the 22 pilots to be robust and generalised. Now concentrate – here comes the science bit!

When fatigue, alertness and sleep were subjectively assessed, in the majority of cases the pilots felt moderately to extremely tired before going to sleep, and much less tired after sleeping, and they evaluated their quality of sleep to be good or average. Interestingly there was a very high degree of correlation between the pilots' self-assessed subjective sleep durations and objective sleep duration obtained from the actiwatches, showing that when it came to assessing sleep the pilots were good judges.

Based upon the baseline performance of the mean PVT value for the first duty day in Block 1, analysis indicated no obvious patterns between the average daily performance and the sleep duration or sleep efficiency. The relationship between subjective alertness measurements with the performance values per sector indicated that in general:

i) better performances (i.e. PVT values) are more often associated with good alertness levels, whilst
ii) the worst performances are more associated with worst alertness levels.

The effect of duty transitions is important as it can be expected that there will be a level of performance decrement associated with duty transitions due to the sustained period of wakefulness and acclimatisation to early-duty sequences operated beforehand. Transitions were defined, in this study, as the changes between the early and late duty days, and rest and duty days.

The effect of transitions on daily average cognitive performance, average alertness level and objective sleep duration indicated that in general the performance on the first late duty day is no worse than that on the previous early duty day and also that performance on the first duty day is no worse than that of the previous rest day. And when it comes to alertness, there is no correlation between early – late days or indeed between rest and duty days. These results indicate no obvious performance decrements associated with duty transitions under the specified FRMS transition guidelines.

During the duty blocks, when it came to self-assessed alertness, the pilots rated themselves as moderately tired. They appear to fully recover after the first duty block to being a little tired. Following the second block and two rest days, the pilots indicated they did not fully recover to this level. This trend was supported by the group PVT variation per block, though the performance variation was slight. A review of cumulative sleep debt per operated block set against a benchmark figure of 8 hours again showed that crew are almost fully recovered after three rest days (disregarding off-duty social interests) and less recovered after two rest days. Based upon these results, easyJet adapted their schedule to ensure adequate rest and recovery for the pilots.