A FRAMEWORK FOR THE ANALYSIS OF THE WORKLOAD OF TRAINING CAPTAINS: A SHORT-HAUL AIRLINE CASE STUDY

Dr. Arnab Majumdar[1]
Mr. Vincent Wu
Dr. Branka Subotic

Centre for Transport Studies
Department of Civil and Environmental Engineering
Imperial College London
South Kensington
London SW7 2AZ
U.K.
a.majumdar@imperial.ac.uk
Phone: (+44) 207 5946037

Captain Simon Stewart
easyJet Airline Company Ltd
Hangar 89, Luton Airport
Bedfordshire LU2 9PF
U.K.
Simon.Stewart@easyJet.com
Phone: (+44) 01582 525 602

Dr. Alexandra Holmes
Clockwork Research Ltd.
21 Southwick Mews
London W2 1JG
U.K.
alex@clockworkresearch.com
Phone: (+44) 207 402 6233

Word Count: 7,307 (includes abstract)

[1] *(corresponding author)*, Email: a.majumdar@imperial.ac.uk

ABSTRACT

Low cost carrier (LCC) airlines that fly short-haul and medium-haul routes have grown rapidly in Europe in recent years. LCCs achieve a low-cost structure by maximising aircraft utilization and streamlining crew scheduling; for example, by flying one type of aircraft and limiting overnight stops. A potential corollary of intensive crew utilisation is elevated levels of crew workload and fatigue, which if not managed effectively, can increase the likelihood of adverse safety outcomes. Training Captains are a sub-group of the pilot population for whom workload may be particularly problematic. These pilots fly standard duties as well as commanding a range of training and assessment duties. To our knowledge, this is the first paper to present a methodology to analyse the factors affecting Training Captains workload with the objective of assisting airlines to manage workload more efficiently. For 15 duties, 29 Training Captains from a LCC completed workbooks that collected information on duty type, work schedule, personal characteristics and sleep, together with perceived workload and alertness ratings. The linear mixed model was used analyse the data. The main finding was that after controlling for scheduling variables, personal characteristics and sleep duration, workload varied systematically across the different duty types. Training CaptainsTraining in the simulator and training whilst flying posed a higher workload strain than standard flying duties, with simulator training duties associated with the highest workload. The results are used to make recommendations on how workload can be managed in an airline.

1. Introduction

Low cost carrier (LCC) airlines flying short-haul and medium-haul routes have grown rapidly in Europe since the introduction of the European Commission's aviation liberalisation packages in 1987. The LCCs business model realises cost savings through measures including "no frills" cabin service and optimal aircraft and crew utilization. Resource utilisation is optimised by operating a singletype fleet out of fixed home bases. Crew work locally and return to their home base at the end of almost every duty, thereby saving on overnight stops and crew transportation. Fairly regular and uniform sectors (point to point) of around 1.5 hours are operated. Crew typically work 2 to 6 sectors and between 8 and 12 hours per duty day.

Where crew utilisation is not managed effectively, it has the potential to bring about elevated levels of crew fatigue and workload. Fatigue has been attributed as a factor in 4 to 7% of serious aviation accidents and has been identified as a significant risk to aviation safety [1]. The most common formal control for crew fatigue involves limiting the number of hours crew can operate. For example, in the UK, the Civil Aviation Authority (CAA) sets out its flight time limitations (FTL) in Civil Aviation Publication (CAP) 371 [2]. While limitations on work hours are a useful safety tool, they have been criticised on the basis that they restrict the number of hours that can be worked without due consideration of the level of workload that is involved in a task.

easyJet is the UK's first LCC and now one of Europe's most successful airlines with nearly a quarter of the LCC market share [3]. The airline recognises the benefits of actively measuring and managing crew workload and fatigue and has conducted a number of research studies. The aim of the current study was to explore the factors that affect the workload of a specialised sub-group of the pilots' population known as 'Training Captains'. These pilots fly standard duties and also conduct the training and examination of other pilots, both in the air and in flight simulators. To our knowledge, the workload experienced by Training Captains (TCs) has not previously been assessed.

This paper presents a methodology for analysing the factors affecting the workload of TCs, using workbooks filled in by pilots for a 3-week period, with the objective of assisting airlines, specifically easyJet, to manage workload more efficiently. This research is timely: as the aviation market continues to grow, with new routes swelling crew numbers and fleet sizes, the role of TCs becomes increasingly more important.

The paper is ordered as follows. Section 2 outlines the duties and role of Training Captains. This is followed by a review of the literature on workload and drivers for pilots in short-haul operations are outlined. This section also outlines the relationship between workload and fatigue with a view to examining the impact of rostering (i.e. scheduling) and duty type factors and sleep characteristics on workload. Section 4 describes the data collection methodology for this study whilst the following section provides an overview of the data and defines the variables for subsequent analysis. Section 6 analyses the data using a general linear model technique to assess significant factors affecting workload, together with a discussion of the results. This is followed by the conclusions.

2. Duties of Training Captains

At easyJet, the route of advancement for aircrew, in ascending order of seniority, is shown in Figure 1. The last three categories are collectively known as Training Captains

From Figure 1, it is apparent that TCs are senior pilots qualified to conduct training and examinations for other pilots. TCs complete standard flying duties, known as line duties In addition, they train other pilots in the line, oftenwhile captaining the flight. Those qualified to train in the simulator complete the aforementioned duties as well as simulator training. The most senior trainers also assess pilots in the simulator. In contrast to standard pilots, TCs spend time outside their duty hours preparing to conduct lessons, and assessments and performing associated administrative tasks. The different types of duty and additional work that the role of training captain involves are expected to add to the workload of standard line flying.

In addition to the above, TCs can be required to conduct "other" duties, such as assessing whether First Officers are ready to commence training to become a Captain (command assessment) and checking performance on the line (line checks). In these cases, the TCs do not instruct or examine pilots in the same manner as for their training duties on the simulator or on the line. Finally, there are duties known as "jump-seat duties" during which TCs observe and train pilots on the line without flying the aircraft at the same time.

Within easyjet, TCs are on average older than line pilots and have flown more hours. The majority of TCs have a commercial flying background, with a small proportion from a military background. Most TCs have had prior simulator training. It is worth mentioning that, when interviewed, pilots with a military background noted that although there was some degree of transferability between military simulator and easyjet simulator training, the differences between the scope and objectives of these two types of simulator training were considerable[2].

Typically, a pilot has flown over 1500 hours for another company prior to joining easyjet as a first officer. A minimum of 3000 flight hours is required before a first officer can become a captain, which may take five years. Once a captain, a minimum of two years is required before a pilot qualifies to become a line-TC.

3. Workload

This section discusses the definition, measurement and causes of workload as relevant to this particular study. It is not intended to provide a detailed review of workload, of which there are numerous examples, such as the European Organization for the Safety of Air Navigation (EOSAN) [4] study.

3.1. The definition of workload

[2] Based upon interviews with easyjet Training Captains conducted on November 6, 2008.

While there is no straight-forward definition of workload in the literature, it is broadly agreed to be a multi-dimensional concept involving various demands placed upon the subject and interactions between a subject and their tasks [5]. It is a construct [6] that cannot be studied directly but can only be inferred from different quantifiable variables [7].

Hilburn [5] made a distinction between taskload i.e., the objective demands of a task, and workload i.e., the subjective demands as experienced by the subject. Factors external to the subject, such as the physical and social work environment, influences how a set of demands translate to taskload. Internal characteristics, such as ability, experience and personality, which together affect the work strategies chosen by individuals, affect how much workload is perceived from a taskload set. Figure 2 illustrates the above ideas in a proposed workload model.

A key point is that workload is in fact the perceived mental strain felt by the individual [7] and different individuals faced with the same task demands under the same conditions can have different levels of workload due to their internal differences, i.e. the cognitive workload is essentially subjective.

3.2. Workload measurement

The numerous workload measurement techniques in the literature can be categorized into three types: performance, physiological and subjective measures. Each measure is based on one or more aspects of the multi-dimensional construct that is workload [8] and there is no universal measure suitable for all aspects under all circumstances. In fact, it is common to find dissociation between measures of different aspects of the same construct [9]. Therefore, Annett [9] has advised that measures should be chosen with the investigation purpose in mind. Furthermore, the lack of standardised workload measures has led the European Organization for the Safety of Air Navigation (EOSAN) [4] to urge researchers to use a battery of techniques if possible.

Subjective measures of workload were selected for use in this study of TC's workload. Subjective measures provide a description of the inner experience of the subject and with it, an indication of the demands on cognitive resources [9]. Given the inherently subjective nature of workload, objective assessments of task demands and performances at best measure objective taskload. There is a stochastic element in moving from taskload to workload, dependent on internal operator characteristics, leading to the intuitive conclusion that subjective workload requires subjective measures.

Another factor in favour of subjective measures is the sheer logistics involved in measuring workload. For example, to undertake physiological measurements of workload in the cockpit after each duty would simply be impractical, as well as pose considerable safety concerns. Furthermore, there is widespread disagreement in the literature regarding the precise nature of the relationship between physiological measurements and workload. Alternatively, whilst it is possible for subject matter experts to rate the workload, this would require their presence on the flight deck. This would not be possible for this study as the training captain often occupies the jump seat, leaving no room for the observer. As Stein and Rosenberg [10] highlight, subjective, self reported measures are more popular in aviation studies due, in part, to their convenience, low cost and safety considerations.

There are numerous subjective measures of workload instigated by the National Aeronautics and Space Administration - Task Load Index (NASA-TLX) and the Subjective Workload Assessment Technique (SWAT) being two of the most widely used [11]. Crucially, both techniques have been used in many aviation studies [9] and, indeed, they were originally created for use in aircrew studies, a fact that augurs well for their construct validity in the Training Captain workload study. For an assessment of the NASA-TLX and SWAT techniques, see [11] and [4]. Both studies rank NASA-TLX as superior to SWAT in crucial aspects including: greater convenience; having been more widely validated, and; having higher sensitivity, especially over low workload ranges.

The NASA-TLX method is a multidimensional scale whereby the overall workload is a function of six dimensions on a continuum. There are two steps to the method: rating and weighting. The rating step requires the subject to rate task(s) according to six dimensions: mental demands, physical demands, temporal demands, effort, performance and frustration level (see Table 3). This feature of NASA-TLX agrees with the multi-dimensional concept of workload; the first three items capture the demands of the task on the subject while the last three items are concerned with the interaction between subject and task. The subject rates each dimension's contribution to perceived workload on a 20-point scale to yield a score out of a hundred and this scale enhances measurement sensitivity. By addressing the individual components of workload, the NASA-TLX method provides diagnostic value in determining workload source.

The second step requires the subject to choose, in each fifteen pair-wise comparisons amongst the six dimensions, the one deemed more important in creating the workload of the task to derive a weight for each dimension. Each dimension's score out of a hundred from the rating step is then weighted accordingly and added up to give a final workload score from 0 to 100. The advantage of weighting is that it gives greater consideration to factors deemed more important in creating workload, further enhancing the sensitivity of the measurement.

The results of the test remain reliable if the test has been administered up to half an hour after the end of the task [13]. This offers the advantage of unobtrusiveness and convenience, alongside those of validty, diagnosticity and sensitivity already outlined. A number of recent studies in aviation have used the NASA-TLX e.g. [7], [14] and [15], demonstrating its extensive application.

3.3. Causes of Pilot Workload

Piloting a modern aircraft for commercial operations is a highly complex task that places great demands on an individual's cognitive capacities [16]. It involves a complex, multidimensional series of behaviours, most of which cannot be observed directly [10]. Thus, pilot workload is an overwhelmingly mental, as opposed to physical, workload.

High workload is a major concern in short-haul flight operations e.g. [17], [18], mainly because of its multi-sector nature with multiple landing and takeoffs. Table 4 summarizes some suggested drivers of aircrew workload (excluding fatigue, which is discussed in the next section), in short-haul flight operations. Two of the key characteristics of the duties of Training Captains; namely, fulfilling a training role and completing non-flying tasks outside a duty, are featured, as are other characteristics of their job; namely the time constraints faced and the number of sectors flown.

3.4. Workload and fatigue

The literature warns about complexity in all parts of the workload-fatigue-performance-safety link in aviation [19, page 310]. Part of the problem is that the workload-fatigue link goes both ways. Various researchers have highlighted workload as a contributor to fatigue in short-haul pilots (summarised in Table 5). For example, [17] found workload and sleep deprivation to be the primary causes of fatigue for short-haul pilots. Similarly, a review of fatigue research [20] reported that for short-haul aircrew, fatigue was most strongly attributed to workload (which it termed 'hassle') and various sleep-related conditions. It is worth noting here that workload is only an important cause of fatigue in short-haul flights and that the causes of fatigue differ between long- and short-haul operations.

While it is well documented that workload is a cause of fatigue, Bourgeois-Bourgine et al. [17] found that fatigue also leads to an increase in perceived workload. To explain this, Cabon et al. [14] postulated that aircrew had to work harder to maintain performance in the face of fatigue In highlighting the link between workload and performance, this explanation adds another layer of complexity. Furthermore, the workload-performance link is also likely to be bi-directional, since greater effort (higher workload strain) is likely to lead to better performance and a bad performance is thought to contribute to higher perceived workload; hence the inclusion of performance as one of the six dimensions of workload in NASA-TLX. With these complex interrelationships in mind, the proposed workload and fatigue model is outlined in Figure 5 and further details can be found in Wu [22].

In the context of this study, aircrew fatigue will be assessed as a function of scheduling [1]. A plethora of research has linked fatigue to scheduling variables, including: consecutive duties; duty timing and duration, and; time of day. The advantage of considering fatigue via scheduling variables is that, compared to a crew's subjective feelings of fatigue and sleep, scheduling variables can be measured and managed far more easily by both airline management and regulators., i.e. they can assesses the impact of scheduling and sleep factors on workload. There is a considerable literature on the impact of duty timing and duration on pilot fatigue[3], but little on the consequences these factors have for workload.

4. Methodology

[3] Fatigue is similar to workload in that it is a construct, difficult to define with precision and with multiple methods of measurement, see Battelle Memorial Institute [23]

A pencil-and-paper workbook was designed to collect the data used in this study. The workbook considered previous aircrew workload studies, easyJet in-house pilot fatigue studies and the advice of a senior TC. Following a review of the draft version of the workbook by six TCs during a three-week trial study, adaptations were incorporated into the final version of the workbook. Figure 4 illustrates this process.

The workbook essentially consisted of two parts. The first part, completed prior to the commencement of their duties, collected information on the personal characteristics of the TCs, such as date of birth and various indicators of experience..

The second part comprised 15 duty assessment forms that were completed at the conclusion of each duty. For each duty, information was collected on the duty type, duration, commute and positioning, relevant sector details, trainee characteristics, sleep details prior to duty and subjective ratings of workload. The perceived workload from each duty was measured using the NASA-TLX method. The weighting component of the NASA-TLX was completed for each duty type.

To reduce post-event memory lapse of the subjective self-rating methods used in this study, the participants were instructed to complete each duty assessment as soon as they finished the duty and definitely before the start of a commute home or to a hotel.

At the beginning of the study period, there were 83 TCs at easyJet. All were approached to participate in the study. Participation in this study was voluntary and participants were assured that all information and results would be de-identified and remain strictly confidential. Twenty-seven TCs were unsuitable for this study for a variety of reasons, e.g. due to leave easyJet soon; working on a part-time basis or flying the line but not engaged in any training duties. Of the remaining 56, valid responses were received from 29 pilots, i.e. 52%.

Finally, this study was conducted from late September to early January. Hence much of the study did not occur during the busy northern summer scheduling season (last Sunday of March till the last Sunday in October each year).

5 Preliminary Analysis

Data was successfully collected for 338 duties. Following data processing, more than a hundred unique variables were derived. It is beyond the scope of this paper to provide summary statistics for every variable, hence, only variables relating to the following items of interest are presented;

- demographic information;
- duty types;
- work schedule and travel time;
- sleep; and
- NASA-TLX ratings.

5.1 Demographic information

The TCs had a mean age of 49 (n = 29, s.d. = 8.03, min = 36, max = 64). They were very experienced having held their ATPL for an average 18.41 (n = 27, s.d. = 7.94, min = 10, max = 39) years and clocking an average of 12,317.59 (n = 27, s.d. = 4,013.857, min = 7 000, max = 24,000) flight hours. The Pearson correlation of these two indicators of experience is very high and significant (0.912, n = 26, 2-tailed significant at 0.01 level) so either should be suitable as an indicator of cockpit experience. Years of holding the ATPL was chosen as the indicator of experience and divided into two categories with 11 subjects having held the ATPL for 15 or fewer years and 16 for more than 15 years.

Ideally, the data on the TCs would be further refined to control for the quality and character of flight experience of each TC by accounting for the overall flight experience of an individual as such information may impact on subjective perceptions and reporting of workload. Such data would include items such as:
i) flight hours in other aircraft aside from the relevant type;
ii) flight hours for other airlines;
iii) flight hours for military training.

Whilst future studies of TC workload should record this information, for this particular study, such an endeavour would have increased the size and scope of the workbook and thereby increased the time for completion by the TCs. In turn, this is likely to have had negative consequences for the reliability and comprehensiveness of data for subsequent analysis.

Age can influence measures of sleep and fatigue. For example, night-time awakening increases with age, resulting in lower sleep quality. Indeed, Yen et al. [18] suggest that fatigue may be pronounced in short-haul pilots over 50 years of age. Consequently, a division by age was made to investigate the effects of age on fatigue, with 50 being the age of interest.

5.2 Duty types

As previously outlined, flying the line is the standard duty carried out by all easyJet flight crew. Following that, line training duties are done by all TCs, while only those with simulator qualifications can conduct simulator duties. Of the 29 trainers, 12 were qualified to undertake line training (TRIs) and 17 were qualified to also undertake simulator training (TRE).

Table 6 gives a detailed breakdown of the different duty types. Similar duty types are grouped into four broad categories for analysis. Hereafter, unless otherwise stated, references to duty types refer to these broad duty categories.

Trainers spent 13% (n=45) of duties on the jump seat during the study duration. The incidence of being on the jump seat was uneven across duty types. As Table 7 indicates, it is far more likely for a trainer to be on the jump seat in a simulator duty than a line duty. It is hypothesized that being in a jump seat instead of at the controls during a duty is associated with lower workload.

5.3 Work schedule and travel time

At the time of the study, the easyJet line pilots were working a fixed-pattern roster consisting of 5 early duties, 2 days off, 5 late duties and 4 days off i.e. 5/2/5/4. The Training Captains were rostered in a much less structured manner and rarely worked a 5/2/5/4 pattern. For example, one training captain worked a sequence of 3/2/4/2/2/2/4/4/3[4], i.e. there was no discernable pattern to his schedule.

Overall, for the 383 duties in total analysed in this study, the majority (73%) of duty blocks consisted of one to three consecutive duties followed by one or more days off.

Duty start times were categorized into 12 two-hour blocks spanning the 24-hour day. The peaks in the distribution at 05:00-06:59 and 11:00-12:59 reflect the two waves of aircraft departures typical of the easyJet operation.

Further investigation revealed that the mean start time for simulator duties was 11:17 (s.d. = 3 hours 49 minutes). The other duty types had mean start times of between 8:30 and 9:30 am.

Rostered (planned) duty duration was considered plus the actual time the TCs took to complete duties. TCs are involved in administrative tasks, such as preparing lesson plans, and the time for these tasks was included in the records of actual duty duration. A final consideration was the travel required prior to commencing and after completion of the duties. This travel takes the form of commute or positioning[5] or both. For this study, time spent positioning on a duty day plus the commute duration comprised the total travel time to the duty. Table 8 shows, for each duty type, the mean rostered duty duration, actual duty duration, time spent completing additional tasks relating to the duty and the sum of the actual duty duration plus the time spent on additional tasks and spent travelling to/from duties.

The data shown in Table 8 indicates that compared to flying the line or other duties, simulator duties are approximately one hour shorter and line-training duties around 40 minutes longer. The discrepancy between rostered and actual duty durations was negligible for the controlled environment in which simulator training takes place. In contrast, when working in the line, operational hassles and delays tended to extend the duty day by approximately 30 minutes. As expected, simulator duties were associated with the highest number of additional tasks (mean=48 minutes). The amount of additional time spent when training and flying the line was comparable. The average travel times for all four types of duty ranged in a narrow band form 49 minutes (for other duties) to 62 minutes (for simulator duties),

For the subsequent analysis:
i) Actual duty duration is divided into three categories: 6 hours or less (36 duties), 6 to 10 hours (228 duties) and more than 10 hours (118 duties);
ii) total travel time is divided into three categories: 30 minutes or less (87 duties), between 30 minutes to one 1 hour (218 duties) and more than 1 hour (75 duties).

5.4 Sleep

---

[4] 3 duties, 2 off, 4 duties, 2 off, 2 duties, 2 off, 4 duties, 4 off and 3 duties

[5] . Positioning is the ‘practice of transferring crew from place to place as passengers in surface or air transport at the behest of an operator.’ [2, page 2]

The quantity of sleep obtained is a fundamental determinant of fatigue. Total sleep duration prior to a duty is the sum of sleep on the previous night and any naps indicated by the subjects prior to duty. A mean sleep duration of 7 hr 14 min (n = 381, s.d. = 1 hr 8 min, min = 4 hr 5 min, max = 11 hr) was obtained during the study.

For subsequent analysis, sleep duration is divided into three categories: less than 6 hours (39 duties); 6 to 8 hours (232 duties) and 8 or more hours (110 duties).

5.5 NASA-TLX ratings

Table 9 shows the weighted component and total NASA-TLX scores by duty type. It can be seen that the dimensions of physical demand and frustration are not significant contributors to workload in most cases. Effort is significant in some cases but contributes less to perceived workload than mental demand, temporal demand and performance. The diagnostics provided by the six components provide evidence for some features of workload mentioned in section 3:

- Workload in aviation is primarily mental, not physical;
- Temporal demand is a significant contributor to workload, possibly due to the demands of rapid turnarounds; and
- Performance is a significant contributor to workload, possibly because pilots feel the need to marshal spare capacity when they perceive their performance to be poor.

Further examination of Table 9 indicates that the average weighted temporal contribution to total workload is highest for simulator duties compared to other duty types. While there are no rapid turnarounds in simulator duties, temporal pressures probably arise from the limited time allowed on simulators for duty completion. Mean mental demand and performance contributions are also highest for simulator duties leading to the highest average TLX scores for simulator duties amongst all duty types.

The distribution of the average NASA-TLX was tested and assumed to be normal for the subsequent analysis.

6. Analysis

A review of relevant studies revealed insights into techniques by which to analyse the data collected in this study. In particular, the diary study of air traffic controllers in Spencer et al. [24], which collected personal information of each subject and details of each duty over a 20-day period, provided useful guidance. Spencer et al. [24] used a form of unbalanced analysis of variance (ANOVA), utilising a two-stage incomplete block analysis; i.e. a mixed model with fixed treatment effects and random block effects [25].

Data from this study was analysed similarly using the linear mixed model, implemented on SPSS 13.0 [26]. The general form of the mixed model is [26 page 7]:

$$\mathbf{y} = \mathbf{X}\hat{\mathbf{a}} + \mathbf{Z}\tilde{\mathbf{a}} + \mathring{\mathbf{a}} \qquad (1)$$

where $\mathbf{y}$ is a vector of responses, $\mathbf{X}$ is the fixed-effects design matrix, $\hat{\mathbf{a}}$ is a vector of fixed-effects parameters, $\mathbf{Z}$ is the design matrix of random effects, $\tilde{\mathbf{a}}$ is a vector of random effects parameters and $\mathring{\mathbf{a}}$ is a vector of residual errors.

Beside fixed and random effects, repeated effects were included in the model in acknowledgement of the fact that the observations in the data set come from only 29 independent subjects so that the assumptions of identical independently distributed errors are unlikely to be valid. The incorporation of repeated effects allows for correlation and non-constant variance of residuals in the mixed model.

In addition to the assumption of normality of the NASA-TLX scores, other necessary assumptions for the statistical model include additivity – that effects of variables in the model are additive -- and linearity in parameters.

The choice of variables analysed was based upon both the literature review and the opinions of the subject matter experts. Ten fixed effects variables were included: duty type, jump seat, number of consecutive days worked, duty start time, duty duration, total sleep duration prior to duty, duration of travel to duty, age and years of holding the ATPL as an indicator of experience. All continuous variables were divided into categories as described in the previous section. Repeated effects were stipulated such that subjects are independent but the duties of each subject are correlated. The random effects model was constructed for each subject to have different intercept terms. However, the smaller model without random effects was judged superior based on the –2 Restricted Log Likelihood information criteria so the final model included fixed and repeated effects.

All variables of n (where n ≥ 2) levels had n-1 parameters estimated for them. The coefficient of each of these n-1 levels is interpreted as the effect on workload scores above or below the reference level in each variable. Table 10 shows the number of levels and parameters of the fixed effects variables. Each variable fitted in the presence of the other variable has a *ceteris paribus* interpretation.

6.2 Results

Parameters that are significant, at least at the 0.05 level, are listed in Table 11. Variables that did not have at least one significant variable include: consecutive days worked, travel duration to duty and experience. Consecutive days worked may not have been significant because as much as 50% of duty blocks consisted of only one or two duties. Similarly, experience may not have influenced perceived workload because the training captain population was very experienced, at least at standard line duties. It is a reasonable hypothesis that more experienced operators have greater success in formulating efficient strategies to meet task demands leading to lower workload. However, the study population may have been past the threshold beyond which additional experience has little effect on workload; for instruction/examination duties (line, simulator and other), it was thought that the workload depends mostly on the innate abilities of individuals and quality of trainees, hence, experience has little effect.[6]

The interpretation of the intercept term is that the average workload of an 'other' duty type where the subject is not on the jump seat, starting between the hours of 2300–0059, lasting more than 10 hours, done by a 50 year or older subject who had 8 or more hours of sleep prior to the duty is 68.20 on the NASA-TLX measure.

Simulator duties account for an increase of 18.86 in workload score over the reference duty type 'other'. Line training duties give a higher workload of 4.39 on average while flying the line is associated with an average of 5.62 lower workload score than the reference duty type. This means that the additional duty types done by Training Captains all impose higher workload strain than flying the line, with simulator duties being the most straining. In addition, after controlling for duty type, being on the jump seat results in lower workload on average.

The importance of temporal pressures for simulator training has already been outlined as a possible cause for the high workload. In addition, it may be that the actual training programme utilised in the simulator differs from that when flying the line. Given the controlled world of simulator training, it may be that the training programme designed for the simulator places greater workload strains on the Training Captains.

Duties between the evening hours of 1700-1859 are associated with higher workload than the reference midnight hours of 2300-0059; the parameters for all other levels are not statistically significant. If duty start time was affecting workload levels through the effects of fatigue, one would expect to see significant parameters in the other levels, particularly the early-start hours as indicated by the literature. Therefore the lack of significance on all other levels could mean that the parameter on the level 1700-1859 was capturing workload effects directly. It may be possible that the higher workload associated with duties starting in the evening is partially a reflection of cockpit duties during periods of heavy traffic over European skies.

As expected, there is positive correlation between duty duration and workload, i.e. shorter duty durations are associated with lower workload ratings.

It was expected that lower sleep duration should result in higher fatigue levels and thus higher perceived workload during duties. However, compared to sleep duration of 8 or more hours, sleep duration of less than 6 hours was associated with a comparable level of workload and sleep duration of 6-8 hours was associated with lower workload. A likely explanation for this finding is that workload is not simply the sleep obtained the night before a duty, but is influenced by sleep history; that is, the sleep obtained over multiple preceding days. Alternatively, higher workload may be perceived from duties following sleep of less than 6 hours through fatigue effects. In addition, if, as part of their personal fatigue management strategies, subjects try to get more sleep prior to what they perceive to be demanding duties, sleep duration of 8 or more hours may be an imperfect proxy for high workload duties.

[6] According to subject matter expert from easyJet

Finally, Training Captains who are younger than 50 have lower perceived workload on average, all other variables considered, than Training Captains who are 50 or more years old. Age is a controversial issue for airline pilots, as witnessed by the debates on the revision of the "Age 60" rule for airline pilots in the USA [27]. The results from this study can be interpreted in a number of ways. It could be that the higher workload of the older pilots is due lack of sleep resulting in greater fatigue and workload. Alternatively, it could be that older pilots have reduced cognitive resources compared to younger pilots. Balancing these factors, the greater experience of older pilots should mean that they have participated in simulator training to a much greater extent than younger pilots and, hence, such training should not impact on their workload rating to the extent it may for younger pilots. Clearly, however, further information is required to provide definitive answers to the age question, including more detailed objective information on sleep, as opposed to self-reported sleep.

7. Conclusions

This paper has outlined a framework by which to estimate the workload of Training Captains in the LCC, easyJet. Training Captains require special attention compared to pilots who fly the line as they are involved in training and examination duties in addition to flying.

The data for this study was collected at a period when the Training Captains at easyJet were concerned that they faced significant pressures to fulfil the training syllabus to the quality required. In particular, they felt that they were preparing more then previously and conducting training in a reduced time frame, whilst fulfilling rigourous training standards. This study used workbooks completed by Training Captains over a 15-day period to assess the impact of various duty, rostering and sleep variables on the perceived workload of the Training Captains. The results are summarised below in Table 12.

Perhaps the most surprising finding is that simulator training imposes the highest perceived workload. Possible explanations for this have been explored earlier, e.g. the time pressure and nature of training in the simulator.

easyJet has a renowned human factors monitoring programme, which has considered the implications of fatigue on safety in considerable detail. A finding from this study is that incorporation of workload management, either as part of fatigue management or as a separate entity, should be considered. In order to do this, the airline should focus upon the different types of training duties used, in particular simulator training and the time pressures associated with it. Furthermore, duty schedules are also crucial, with the airline managing schedules for pilots such that they do not perceive high workloads from duty start times and durations.

1. Caldwell, J. A. (2005). Fatigue in aviation. Travel Medicine and Infectious Disease, 3, 85-96
2. CAA (2004). *The Avoidance of Fatigue in Aircrews, Guide to requirements*. CAP371. The Stationary Office, Norwich.
3. ECA (2006). *Upheaval In The European Skies – Low Cost Carriers in Europe: Economic Data, Market and Pilot Demand Forecast*, 2nd ed. European Cockpit Association, Brussels.
4. European Organization for the Safety of Air Navigation (2003). *Review of workload measurement, analysis and interpretation methods*. CARE-Integra-TRS-130-02-WP2-1-0. [Online] Available from: http://www.eurocontrol.int/integra/gallery/content/public/documents/expt_def_wp2_final.pdf#search=%22CARE-Integra-TRS-130-02-WP2-1-0%22
5. Hilburn. B. (2004). Cognitive complexity in air traffic control: a literature review. Eurocontrol Experimental Centre [online] Available from: http://www.eurocontrol.int/eec/gallery/content/public/documents/EEC_notes/2004/EEC_note_2004_04.pdf
6. Weinger, M. B., Reddy, S., B., & Slagle, J. M. (2004). Multiple Measures of Anesthesia Workload During Teaching and Nonteaching Cases. *Anesthesia and Analgesia*; **98**, 1419-1425.
7. Averty, P., Colet, C., Dittmar, A., Athenes, S. & Vernet-Maury, E. (2004). Mental Workload in Air Traffic Control: An Index Constructed from Field Tests. *Aviation, Space, and Environmental Medicine*, **75** (4), 333-341.
8. Baldwin, C.L. & Coyne, J.T. (2003). *Mental Workload as a Function of Traffic Density: Comparison of Physiological, Behavioral, and Subjective Indices*. Proceedings of the 2nd International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, 21st-24th July, 2003. Utah, USA.
9. Annett, J. (2002). Subjective rating scales: science or art? *Ergonomics*, **45** (14), 966-987.
10. Stein, E. S. & Rosenberg, B. L. (1983). *The Measurement of Pilot Workload*. Federal Aviation Administration Technical Center. Interim Report number: DOT/FAA/EM-81/14
11. Luximon, A. & Goonetilleke, R. S. (2001). Simplified Subjective Workload Assessment Technique. *Ergonomics*, **44** (3), 229-243.
12. NASA Human Performance Research Group (1987). NASA Task Load Index (TLX) V 1.0 Users Manual. NASA Ames Research Centre, Califonia.
13. Rubio, S., Diaz, E., Martin, J. & Puente, J.M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology: An International Review*, **53**(1), 61-86.
14. Cabon, P., Bourgeois-Bougrine, S., Mollard, R., Coblentz, A. & Speyer, J.J. (2000). Fatigue of short-haul flights aircrews in civil aviation: Effects of work schedules. In Hornberger, S., Knauth, P., Costa, G. & Folkard, S. (ed.), *Shiftwork in the 21st century: challenges for research and practice*. Peter Lang, Frankfurt, pp. 79-85
15. Sohn, S.Y. & Jo, Y.K. (2003). A study on the student pilot's mental workload due to personality types of instructor and student. *Ergonomics*, **46**(15), 1566-1577

16. Wilson, G. F. (2001). An Analysis of Mental Workload in Pilots During Flight Using Multiple Psychophysiological Measures. *The International Journal of Aviation Psychology*, **12**(1), 3-18.
17. Bourgeois-Bougrine, S., Carbon, P., Gounelle, C., Mollard, R. & Coblentz, A. (2003). Perceived Fatigue for Short- and Long-Haul Flights: A survey of 739 Airline Pilots. *Aviation, Space, and Environmental Medicine*, **74** (10), 1072-1077.
18. Yen, J-R., Hsu, C-C., Ho, H., Lin, F-F. & Yu, S.-H. (2005). Identifying flight fatigue factors: An econometric modeling approach. *Journal of Air Transport Management*, **11**, 408-416.
19. Goode, J. H. (2003). Are pilots at risk of accidents due to fatigue? *Journal of Safety Research*, **34**, 309-313
20. CAA (2005). *Aircrew Fatigue: A Review of Research Undertaken on Behalf of the UK Civil Aviation Authority*. The Stationary Office, Norwich.
21. Bennett, S. A. (2003). Flight crew stress and fatigue in low-cost commercial air operations – an appraisal. *International Journal of Risk Assessment and Management*, **4**(2/3), 207-231.
22. Wu, V. (2006) A statistical analysis of training captain workload: a short-haul airline case study**,** MSc. Dissertation, Imperial College London, Unpublished.
23. Battelle Memorial Institute (1998). *An Overview of the Scientific Literature Concerning Fatigue, Sleep, and the Circadian Cycle*. Prepared for the Office of the Chief Scientific and Technical Advisor for Human Factors, Federal Aviation Administration.
24. Spencer, M.B., Rogers, A.S., Birch, C.L. & Belyavin, A.J. (2000). A diary study of fatigue in air traffic controllers during a period of high workload. In Hornberger, S., Knauth, P., Costa, G. & Folkard, S. (ed), *Shiftwork in the 21st century: challenges for research and practice*. Peter Lang, Frankfurt, pp. 251-256.
25. Kendall, M.G. & Stuart, A. (1976). The Advanced Theory of Statistics, Vol. III: Design and Analysis, and Time-series, 3rd ed. Charles Griffin & Company Limited, London, pp. 149-155.
26. SPSS Inc. (2005). *Linear Mixed-Effects Modeling in SPSS: An Introduction to the MIXED Procedure*. [Online] Available from: http://www.spss.ch/upload/1126184451_Linear%20Mixed%20Effects%20Modeling%20in%20SPSS.pdf
27. Turley, J (2008) Letter to The Honorable Nancy Pelosi, Speaker United States House of Representatives. [Online] Available from: http://www.age60rule.com/

.

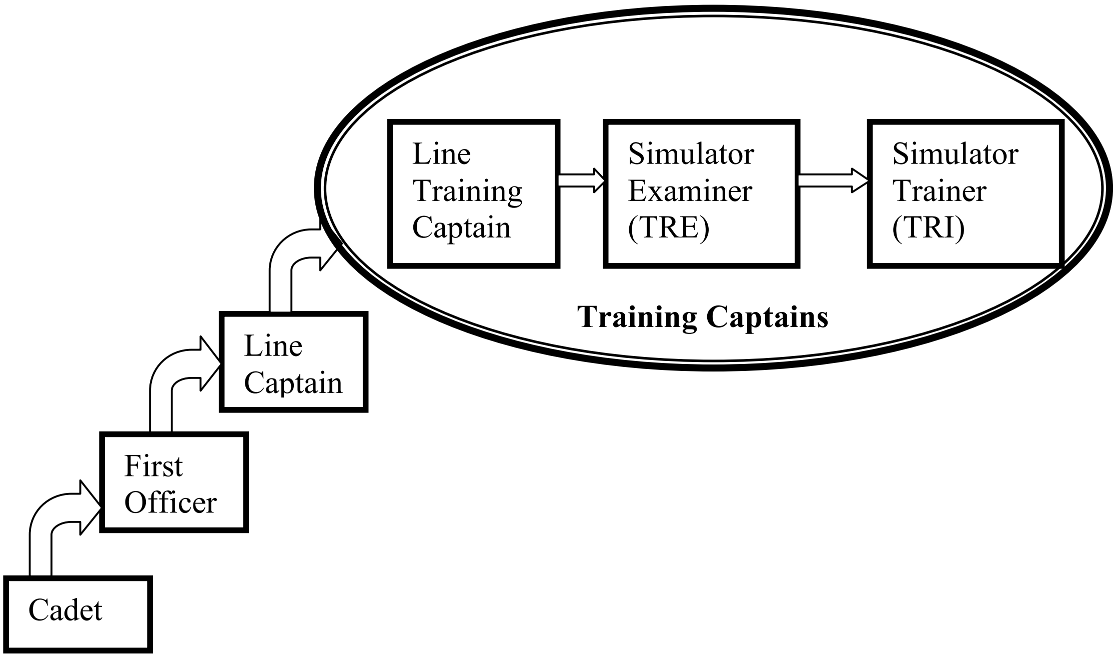

Figure 1. The order of advancement of pilots at easyJet.

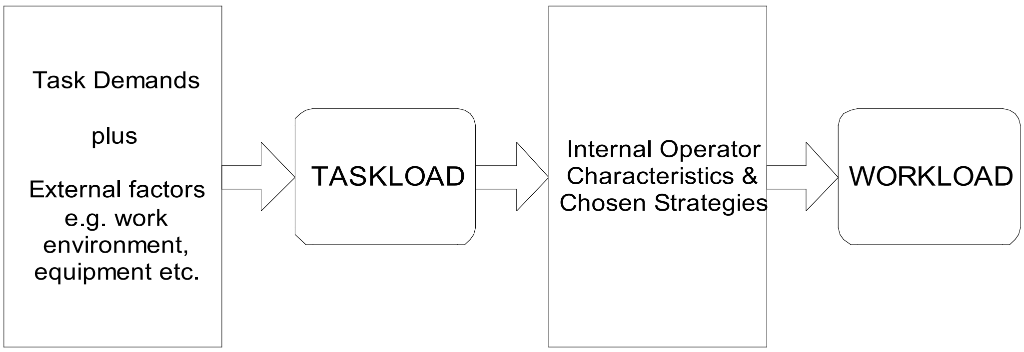

Figure 2: Proposed workload model, Source: Hilburn [5].

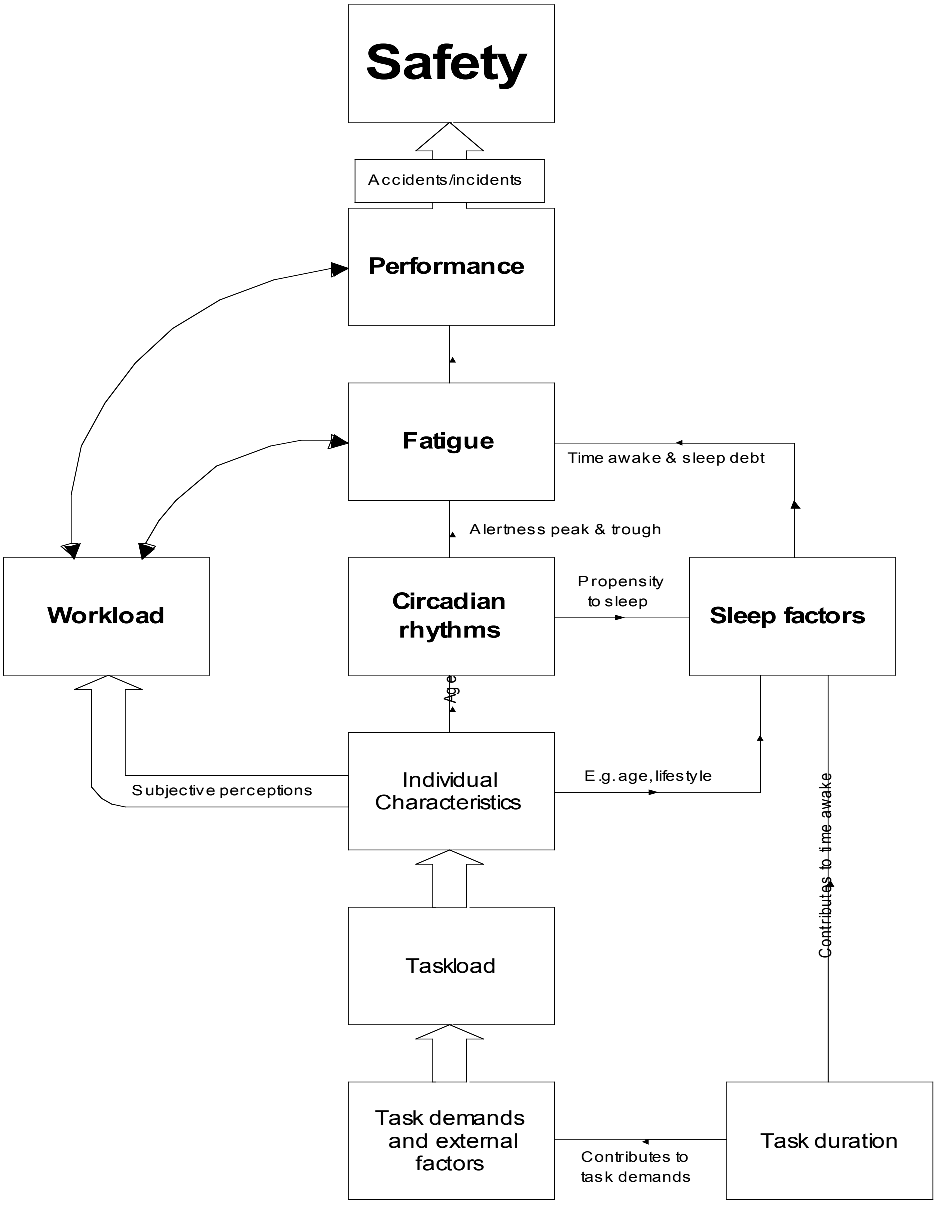

Figure 3: Proposed workload-fatigue model in aviation, Source: Wu (2006) [22]

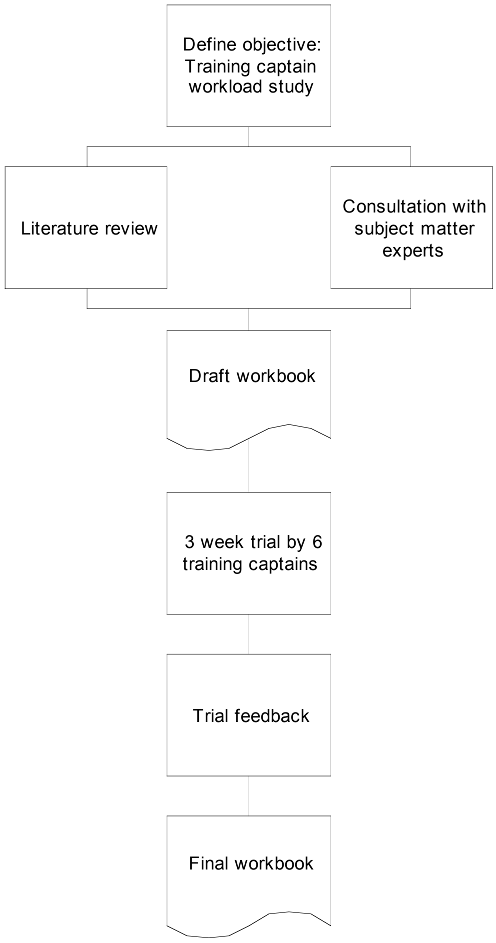

Figure 4. Flowchart of workbook design process

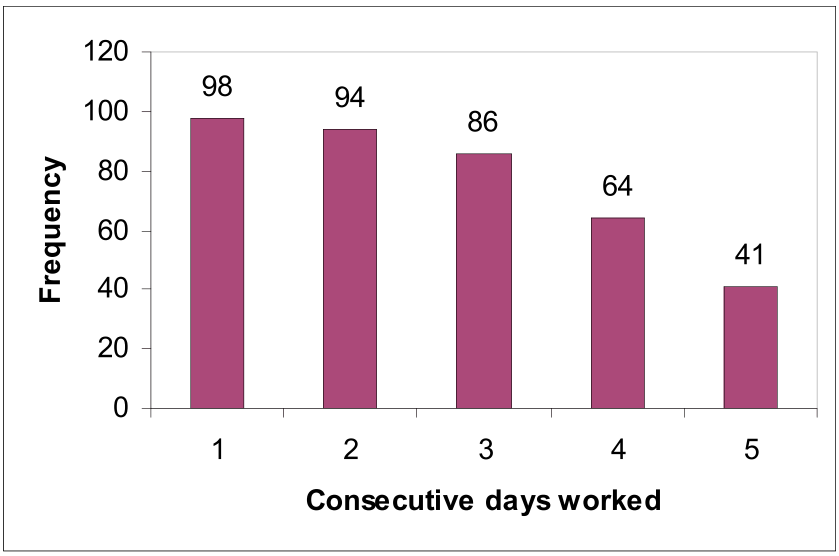

Figure 5: Consecutive days worked frequency

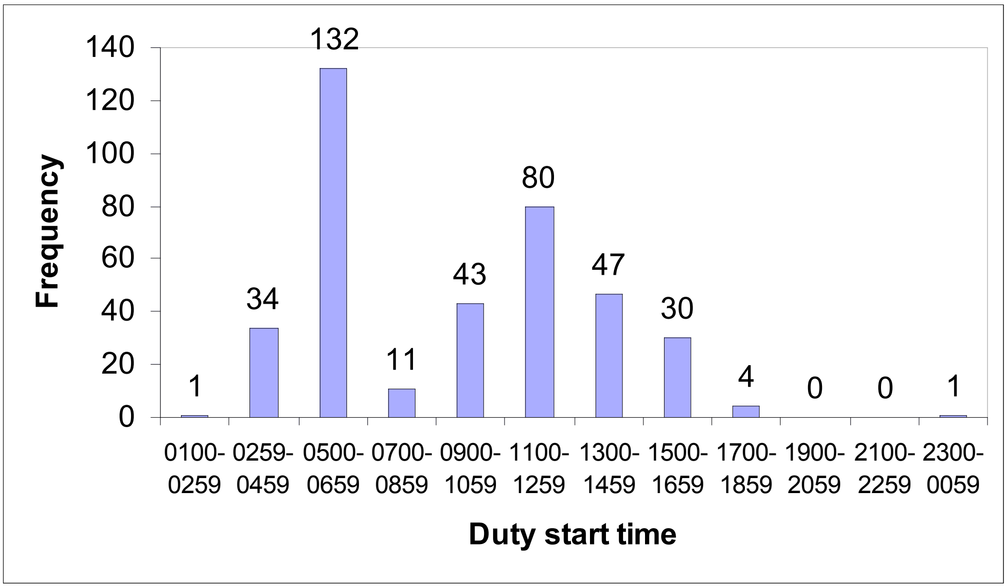

Figure 6: Duty start time distribution

Table 1 outlines the major characteristics of line training and simulator training duties.

Characteristic	**Line Training**	**Simulator Training**
Typical number of trainees	One	Two
Instruction or examination	Instruction only	Both
On the Jump Seat	Rarely	Sometimes
Rostered Duty Start Time	Early or late mornings	Mostly late mornings
Rostered Duty Duration	> 6 hours	6 hours

Table 1: Characteristic of duties of Training Captains for line and simulator training.

Table 2 provides additional information on the types of training conducted.

Training type	**Frequency of training**	**Place of training**
Recurrent	Once every 3 months for all pilots.	Simulator only
Line	On-going, regular training of pilots.	Line only
Refresher	Whenever a specific training is identified for a pilot.	Simulator and on the line
Command	Whenever a first officer is groomed to become a captain.	Simulator and on the line
Zero sector time	Whenever a pilot is changing type of aircraft flown.	Simulator and on the line

Table 2: Characteristics of the types of training that Training Captains have to undertake.

TITLE	ENDPOINTS	DESCRIPTIONS
MENTAL DEMAND	Low/High	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	Low/High	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	Low/High	How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
PERFORMANCE	Good/Poor	How successful do you think you were in accomplishing the goals of the task set by the experimenter? How satisfied were you with your performance in accomplishing these goals?
EFFORT	Low/High	How hard did you have to work (mentally and physically) to accomplish your level of performance?
FRUSTRATION LEVEL	Low/High	How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

Table 3: Rating scale definitions and endpoints of NASA-TLX
Source: Rubio et al. [12]

Elements	Source
Additional non-flying tasks outside duty	Yen et al (2005) [18]
Communicating in a foreign language	Bourgeois-Bougrine et al (2003) [17]
Coordination/personality issues with crew/trainee	Bourgeois-Bougrine et al (2003) [17] Sohn & Jo (2003) [15]
Delays	Bourgeois-Bougrine et al (2003) [17] CAA (2005) [20]
Density of verbal exchange with Air Traffic Control	Bourgeois-Bougrine et al (2003) [17]
Difficult flight	Bourgeois-Bougrine et al (2003) [17]
Interruption during activities	Bourgeois-Bougrine et al (2003) [17]
Number of sectors	Bennett (2003) [21] Bourgeois-Bougrine et al (2003) [17] CAA (2005) [20] Yen et al (2005) [18]
Simultaneous actions	Bourgeois-Bougrine et al (2003) [17]
Technical fault	Bennett (2003) [21]
Time constraints	Bennett (2003) [21] Bourgeois-Bougrine et al (2003) [17] CAA (2005) [20] Yen et al (2005) [18]
Training role during flight	Bennett (2003) [21]

Table 4: Causes of workload in short-haul operation

Element	Cause of fatigue	Source
Early starts	Sleep (time since awake) Circadian	Bourgeois-Bougrine et al (2003) [17] CAA (2005) [20] Caldwell (2005) [1]
Consecutive duty days	Sleep (sleep debt)	Bourgeois-Bougrine et al (2003) [17] CAA (2005) [20]
Long duty day/duty period extension	Sleep (time since awake)	Bourgeois-Bougrine et al (2003) [17] Caldwell (2005) [1] Goode (2003) [19]
Duty time of day (e.g. nocturnal duty)	Circadian	Bourgeois-Bougrine et al (2003) [17] CAA (2005) [20] Caldwell (2005) [1]
Adapting between series of early and late duties	Circadian	Bennett (2003) [21]

Unsatisfactory hotel arrangements	Sleep (sleep debt)	Bennett (2003) [21] CAA (2005) [20]
Lengthy commute	Sleep (time since awake)	Bennett (2003) [21] CAA (2005) [20]
Disturbed sleep on night before	Sleep (sleep debt)	Yen et al (2005) [18]
(see Table 2.1)	Workload	(Table 2.1)

Table 5: Causes of fatigue in short-haul operations

Duty Category		n	%
Broad	Specific		
Simulator	SIM refresher training	2	0.5
	SIM recurrent training	44	11.5
	SIM command training	9	2.3
	SIM zero sector time	3	0.8
	Sub-total	58	15.1
Line training	Line refresher training	5	1.3
	Line training	112	29.2
	Line command training	14	3.7
	Line zero sector	19	5.0
	Sub-total	150	39.2
Flying the line	Flying the line	142	37.1
	Sub-total	142	37.1
Other	Command assessment	13	3.4
	Line check	20	5.2
	Sub-total	33	8.6
	Total	383	100.0

Table 6: Duty types

Jump seat	Simulator	Line	Flying the line	Other	Total
Yes	28	8	0	9	45
No	21	129	121	22	293

Table 7: Jump seat status by duty type

Variable	Duration of duty in hours and minutes			
	Simulator	Line Training	Flying the line	Other
Rostered duration	7:11 (01:33)	8:52 (02:02)	8:10 (01:57)	8:11 (02:01)
Actual duration	7:16 (01:27)	9:15 (02:00)	8:29 (02:09)	8:47 (02:08)
Additional tasks	0:48 (01:22)	0:29 (01:26)	0:23 (01:09)	0:37 (01:31)

Sum of actual duration, additional tasks and travel duration	9:06 (01:32)	10:39 (02:00)	9:54 (02:24)	10:13 (02:07)

Table 8. Average duty duration and travel times, in hours and minutes with standard deviation in brackets.

		Mean (standard deviation)			
		Simulator	**Line**	**Flying**	**Other**
Weighted component	Mental	19.27 (6.21)	16.62 (6.16)	11.76 (6.73)	13.86 (6.66)
	Physical	1.34 (1.80)	1.14 (2.27)	0.92 (2.08)	2.21 (4.76)
	Temporal	16.78 (8.19)	12.68 (7.66)	11.08 (6.30)	12.20 (7.12)
	Performance	16.23 (7.02)	14.88 (6.64)	12.29 (7.90)	13.21 (6.66)
	Effort	11.55 (6.52)	12.84 (7.80)	9.13 (6.77)	11.26 (7.70)
	Frustration	4.75 (5.21)	5.00 (5.33)	6.02 (6.69)	3.75 (4.17)
Total NASA-TLX score		69.88 (15.05)	62.90 (12.62)	51.50 (16.50)	55.96 (15.73)

Table 9: Weighted component and total NASA-TLX scores by duty type

Fixed Effects	**Number of**	
	Levels	**Parameters**
Intercept	1	1
Duty Type	4	3
Jump seat	2	1
Consecutive days worked	5	4
Duty start time	12	11
Duty duration	3	2
Travel duration	3	2
Total sleep duration	3	2
Age	2	1
ATPL years (experience)	2	1

Table 10: Number of levels and parameters of fixed effect variables

Variable (no. of significant /total no. of	**Level**	**Coefficient estimate**	**Std. Error**	**T**

parameters)				
Intercept (1/1)		68.20	4.68	14.57
Duty type (3/3)	Simulator	18.86	1.17	16.07
	Line training	4.39	1.12	3.93
	Flying the line	-5.62	1.09	-5.16
	Others	-	-	-
Jump seat (1/1)	Yes	-8.61	1.13	-7.59
	No	-	-	-
Duty start time (1/8)	17 00 – 18 59	16.51	5.76	2.87
	23 00 – 00 59	-	-	-
Duty duration (2/2)	Up to 6 hours	-4.36	1.08	-4.03
	6 to 10 hours	-1.75	0.86	-2.03
	More than 10 hours	-	-	-
Total sleep duration (1/2)	6 to 8 hours	-1.78	0.82	-2.17
	8 hours or more	-	-	-
Age (1/1)	Less than 50 yrs	-8.19	2.23	-3.67
	50 yrs or more	-	-	-

Table 11: Fixed effect parameters significant at least at the 0.05 level

Independent variable	Findings
Duty type	Simulator duties are associated with the highest workload, followed by line, other and flying the line
Jump seat	Being in a jump seat during duty is associated with lower workload
Duty start time	Duties starting during the period of 1700-1859 are associated with higher workload than duties starting during the reference period of 2300-0059
Duty duration	Longer duty durations are associated with higher workload
Total sleep duration	Sleep duration of 6-8 hrs is associated with lower workload than 8 or more hours
Age	Age 50 and above is associated with higher workload
Consecutive days worked	No statistically significant finding
Travel duration	No statistically significant finding
Experience	No statistically significant finding

Table 12: Summary of findings