Appendix 3. Operationalized management factors for paired comparisons

In this document, management factors have been operationalized to link to each of the base events of the human performance model (HPM) in CATS.

The HPM must be able to represent managerial and organizational influences on human performance. In the causal risk model, safety management is described as providing the resources and criteria for the frontline workforce to operate safely. Resources and criteria are generically described by delivery systems for both hardware and human elements. Human error can be controlled by the following delivery systems

- Procedures, rules and checklists to guide behaviour
- A good man-machine interface. So that it can be operated easily and correctly
- Manpower planning and availability of people to do tasks
- Competence and suitability
- Communication, coordination between online risk controllers
- Commitment, motivation and conflict resolution

Table 3.1: Initial selection of performance shaping factors

Delivery system	Selected PSF	Operationalized PSF	Management influences
Procedure	left out		
Competence	Experience of captain & FO	total number of hours flown	Tab 2
	Training of captain and FO	the number of days since the last type recurrent training	Tab 2
Availability	fatigue	Stanford Sleeping Scale	Paired comparison
	workload	number of times the crew members have to refer to the abnormal/ emergency procedures	Paired comparison
Commitment/motivation	left out		
Communication	Intra-cockpit communication	number of flights in which the pilot and first officer will have a different mother tongue	Tab 2
Man-machine interface	technology interface	four aircraft generations	Tab 2
Weather		Rain fall rate that pilots experienced during the flight	Paired comparison

The first two columns in Table 3.1 present the initial list of selected performance shaping factors together with the associated delivery systems. Some factors which we have not been able to represent in objectively quantifiable units, i.e. procedure and commitment, have been left out. Figure 3.1 shows the structure of the flight crew performance model.

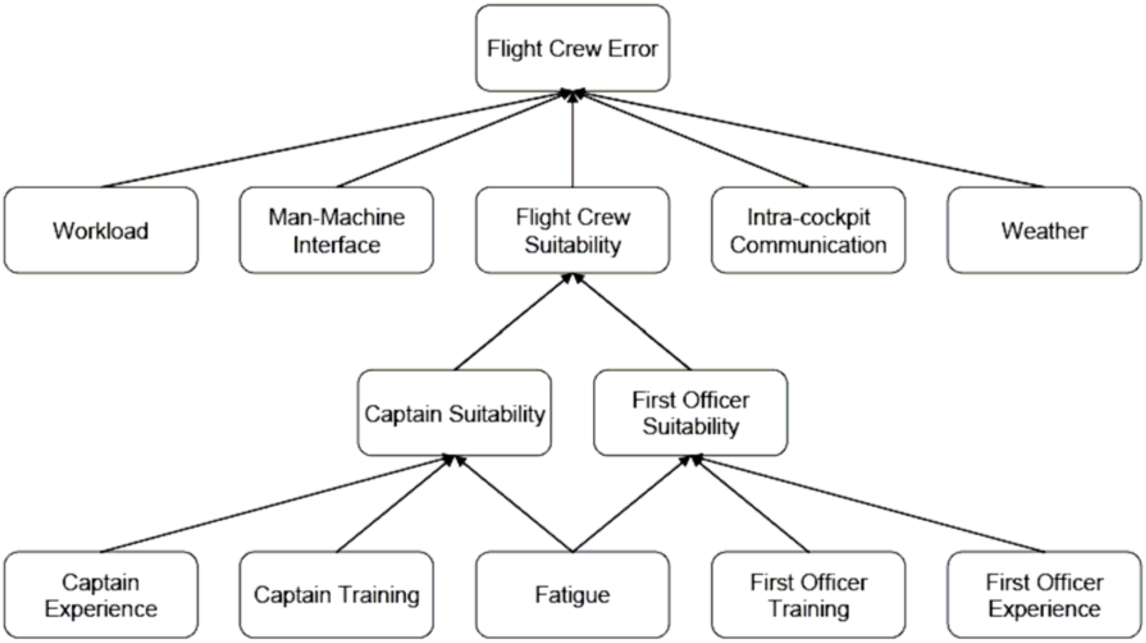


Figure 3.1: HPM model

In the HPM model, the selected PSF have been operationalized as shown in the third column of Table 3.1.

Some delivery system nodes are too complicated to represent at this stage, or we are unable to quantify them in numerical units, e.g. the quality of training. The definitions of the nodes have therefore been severely restricted and modelled in a way that can be easily quantified by the BBN. However, this has potentially biased the original definition of the nodes and lost many important influences. This makes the modeling of the management influences on some of the nodes very limited.

In the following, I have come up with several management items per parameter. In order to define management functions exactly for the nodes defined in CATS. Each definition of the node and the logic behind it are cross referenced here. After internal circulation in the research team the list was presented to one of our expert pilots who was asked to add any management influences he missed and to indicate any he felt were not relevant, using a card sort of the influences per parameter into three groups-very important, marginally important and in between.

AE procedure: number of times the crew members have to refer to the abnormal/ emergency procedures section of the aircraft operation manual during flight

In the context of flight crew activities, workload can be defined as all the physical and mental effort required to fly an aircraft. It includes planning, thinking, navigation, communication, and controlling the aircraft [Stein & Rosenberg 1983]. High workload exists when task demand is close to the operator's maximum capacity, while workload is low when task demand is much below the operator's capacity. Hence, workload is not only sensitive to multiple characteristics of a task, i.e. task demand, but also of the operator, i.e. operator capacity. [Hart 1987, Hancock et al. 1995]. Operator capacity is highly influenced by the fatigue, training and experience of the flight crew and by their communication with each other. This is already represented in other nodes of the flight crew operator model. To avoid double counting we will therefore only consider aspects of task demand. Task demand is determined by the type of action, the number of actions, the sequence of the actions and the time required for the action to be completed [Lysaght et al. 1989].

Factors that influence task demand are traffic density, weather, complexity of the route structure, possible flight delays, amount of information available, and the technical status of the aircraft. An abnormal technical status of the aircraft such that the flight crew has to apply the abnormal or emergency procedures from the Aircraft Operations Manual directly increases task demand. Although each of the mentioned aspects does play a role, it is difficult to clearly define and quantify the combination of these aspects. Arguably the most important factor is the traffic complexity, but as of yet we have not been able to represent this in objectively quantifiable units. The technical status of the aircraft on the other hand is easily defined as whether or not there is a situation that requires the flight crew to consult the abnormal procedures section of the Aircraft Operations Manual (AOM). Therefore this has been selected as the proxy variable for 'time pressure and workload'.

Emergency and abnormal situations are often time critical, complex, and/or ambiguous. The main situations[1] defined in JAR ops when the pilots have to refer to A/E procedure can be roughly categorized into crew incapacities, adverse weather, and technical failure. Although the situations in which the crew have to refer to A/E procedures can be complex, the emergency and abnormal procedures are generally focused on aircraft systems malfunctions rather than on the situation as a whole. Therefore, from a management point of view and in CATS we only focus on situations

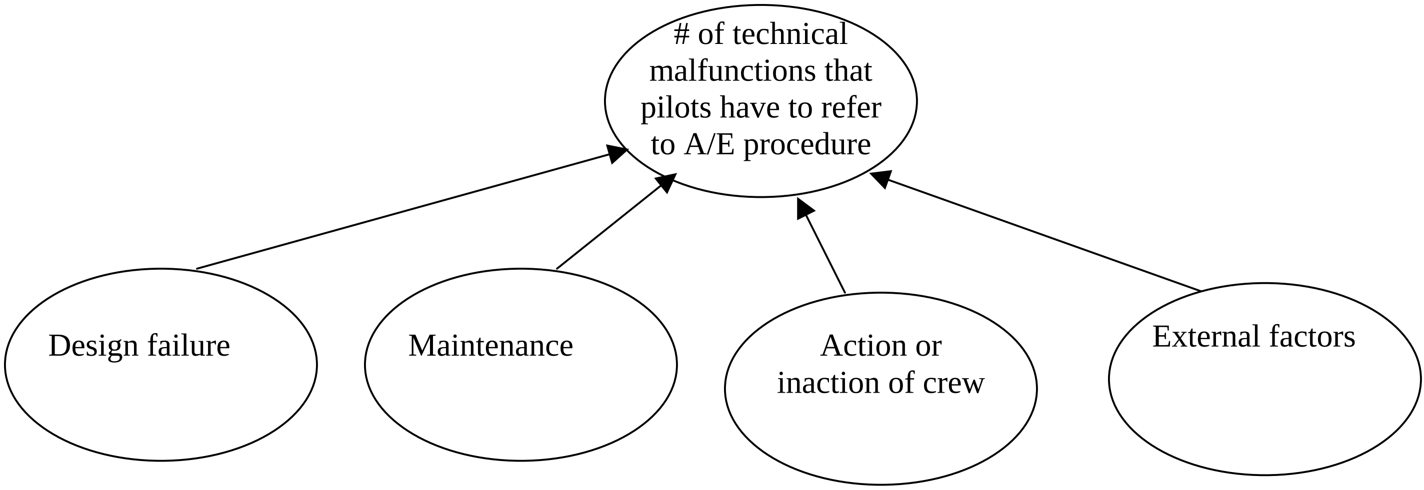
---

[1] In the following situation, abnormal and emergency procedures must be executed (jar-ops from JAA http://www.jaat.eu/publications/jars/jar-ops-1.pdf):
(a) Crew Incapacitation;
(b) Fire and Smoke Drills;
(c) Unpressurised and partially pressurized flight;
(d) Exceeding structural limits such as overweight landing;
(e) Exceeding cosmic radiation limits;
(f) Lightning Strikes;
(g) Distress Communications and alerting ATC to Emergencies;
(h) Engine failure;
(i) System failures;
(j) Guidance for Diversion in case of Serious Technical Failure;
(k) Ground Proximity Warning;
(l) TCAS Warning;
(m) Windshear;
(n) Emergency Landing/Ditching; and
(o) Departure contingency procedures.

arising from technical failures in the aircraft and the management action that can be taken to reduce them.

Technical failures can be classified into 4 categories: inherent technical failure (design), maintenance related, mismanaged by crew, technical failure due to external factors. The first two factors can be dealt with using the design and maintenance processes in the hardware delivery system, hence influences such as suitable intervals for maintenance and replacement, monitoring of technical performance in flight, the minimum list of hardware which has to be present to fly, etc.

To determine the marginal distribution of these four categories where pilots have to refer to the A/E procedure (see figure below), we use structured expert judgment (from pilots) for the probability distribution of 4 variables. If later data is available for us, we will derive information based on data.



Later, when developing the maintenance model, we will design questions to elicit the relative importance of management influences on improving design to reduce failure and optimizing maintenance related issues. For those developments the experts needed will not be pilots, but aircraft designers and maintenance experts.

This node (AE procedure) is a good example to demonstrate that the quantification proposed in CATS only captures a very partial set of the complete influences. We have found that many of the issues concerned, for example, in the International Symposium on Emergency and Abnormal Situations in Aviation Symposium, are not able to be modelled in this CATS management perspective, if this node is operationalized as it is now. Important issues such as follow are not included.

- crew loss of situation awareness for abnormal and emergency situation
- crew lack of knowledge to determine the problem or identify the correct procedure
- the quality of procedures, e.g. the ambiguities in the checklists, checklists not up to date with current regulations, font size, lack of substantive human performance guidelines

- procedure design added to concurrent task demand may reach the crew cognitive limits and increase crew workload
- Intentional non-compliance, e.g. performing several checklists from memory
- Requirements for exceptionally high levels of coordination inside and outside (ATC, maintenance) of the airplane
- A/E procedures are seldom practiced (twice a year or less) and used rarely

Measured physical condition using the Stanford Sleepiness Scale

A person's physical condition has significant influence on performance. The physical conditional is affected by a myriad of factors such as eating habits, age, diseases, the use of medication[2], and the time of day. Generally speaking people perform less well during the early hours in the morning [Rosekind et al 1994]. Representing that influence in a model for flight crew performance is complicated by the fact that pilots regularly cross time zones. Up to 10 time zone crossings in a single flight are no exception. It is therefore rather arbitrary what should be considered 'time of day' for a pilot: local time at the location of the aircraft or local time at the point of departure? And what should be considered the 'time of day' for a pilot who flies from Amsterdam to Los Angeles (crossing 9 time zones), has a stopover of one night, and then flies back to Amsterdam? Another point of consideration is the correlation between time of day and fatigue. Flight crew fatigue has also been the subject of extensive research [Simon & Valk 1993, Simons et al 1994, Simons & Valk 1997, Simons & Valk 1998, Valk & Simons 1996] and data on fatigue levels is available from that research. Because fatigue is less complicated to represent, because of the availability of data, and because time of day and fatigue are correlated, it is proposed to represent only fatigue in the model, and not to represent time of day or any other factor that could influence a pilot's physical condition.

Fatigue will be quantified using the Stanford Sleepiness Scale (SSS). The result of the SSS is a score with increasing sleepiness from 1 to 7, where 1 signifies "feeling active and vital; wide awake" and 7 stands for "almost in reverie; sleep onset soon; losing struggle to remain awake". SSS measures are highly correlated with flying performance and threshold of information processing speed during periods of intense fatigue.

Availability

- Scheduling policy
  - considers 1) entire duration of the flight 2) pre-operating deadhead time 3) training periods prior to a flight 4) office time prior to a flight, to determine rest periods and flight duty time (IOSA)
  - Incorporate flight time accrued in operations other than those of the air carriers in the calculation of flight and duty time (IOSA)
  - Set maximum hours per flight duty period, as determined by the local start time and the number of sectors to be flown (http://www.casa.gov.au/aoc/fatigue/regs/part02.htm)
  - Set maximum hours on cumulative duty period (http://www.casa.gov.au/aoc/fatigue/regs/part02.htm)

- Crew rest
  - Set a minimum rest period according to the duty period and time zone(http://www.casa.gov.au/aoc/fatigue/regs/part02.htm)
  - Set a minimum period free of all duty after a consecutive days of duty (http://www.casa.gov.au/aoc/fatigue/regs/part02.htm)
  - A average sleep requirement for 8 hours in a 24-hour period (Goode, 2003)
  - Provides comfortable accommodation for getting good sleep at stopovers

---

[2] These are important to model in a complete model, but get left out with the operationalisation chosen here

- Creating a suitable nap environment and an appropriate placement of a nap relative to the sleep deprivation period in multicrew aircraft (NASA, 2006 http://human-factors.arc.nasa.gov/zteam/) (Caldwell, 1997)
- Provide several days for the flight crew to adjust to any new sleep/ wake schedule (Caldwell, 1997)

Conflict resolution at management level

- Ensure that management policy is not overridden in practice by over-scheduling tired pilots

Competence-Selection and screening process

- A education and training module that help pilot to understand the cause and effect of fatigue, and teaches pilots how to minimizing circadian rhythm (e.g. use of blight light exposure)
- Forbids alcohol and drug consumption for a suitable period before flying (Caldwell, 1997)
- Screens health and drug/alcohol/medication use before flying

Man-machine interface

- Using good fatigue assessment tools to objectively discern pilots with relatively more fatigue and performance decrement
- Technical alert system that informs pilots who fall asleep during the operation
- Equipment design to improve work condition to reduce operator's on line fatigue and discomfort

Communication, coordination and online supervision delivery system

- A good communication between flight crew members to openly discuss about their current ability to carry on work and rotate flight tasks and converse with other crewmembers

D Language: Number of flights in which the pilot and first officer will have a different mother tongue

The composition of the flight crew is expected to influence flight crew coordination and cooperation and hence affect the flight crew error probability. The ability of captain and first officer to work together as a team depends on cultural and psychosocial (personality) factors and can also be influenced by experience, CRM training, and airline procedures. Trans-cockpit authority gradient is a potentially relevant factor. A steep trans-cockpit authority gradient may exist when the captain has much more experience than the first officer, but it is also influenced by national culture. At this stage it is considered to be too complex to represent trans-cockpit authority in the flight crew performance model, but it is recommended to continue exploring possible ways to represent this aspect in future versions of the model. Intra-cockpit communication is also an essential element of flight crew coordination. Without communication the flight crew cannot work together as a team. It is affected by the native language of the captain and the first officer. For the purpose of this model, flight crew composition and communication is represented by the factor 'delta mother tongue' which represents whether the captain and first officer speak a different native language[3].

Communication is not only important within the cockpit but also from the cockpit to ATC. The aspect of communication between flight crew and ATC will be covered when the flight crew performance model is linked to the other parts of the model, including those parts of the model that represent performance of the air traffic controller, and will not be discussed in this report because at the time of writing the air traffic controller model was still under development.

This node is relatively difficult for management to influence because mother tongue is a very fixed characteristic of a pilot. We can influence this node either by not scheduling pilots of different mother tongue to operate together, or by hiring only pilots who speak the same language. We do not need any expert judgement to estimate the effects of these policies, since we already have data about how the distribution of 'delta language' is for the current situation and what effect this has on errors. To see what implementing either policy would be, we simply have to make 'delta language' = 0

Availability

- Ensure scheduling and manpower planning to prevent pilot flight crew members of different mother tongue from operating together

Competence

- Employ pilots only with the same mother tongue

---

[3] Another possibility would have been to judge competence and training in a common language (usually English)

Weather: the rainfall rate in mm/hr that flight crew experienced during a flight

Weather influences aircraft safety in a complex way. Some weather phenomena are a direct hazard in itself, for instance lightning, microbursts or conditions that result in airframe icing. Other conditions may influence the degree of difficulty of the flying task, such as gusting winds or reduced visibility due to heavy rain or fog. Weather may also create a less comfortable working environment, for example prolonged flight in turbulence is physically rather demanding. Arbitrarily, the rainfall rate has been selected as the indicator of weather in this model. The rainfall rate determines the colour painted by the weather radar in the cockpit and therefore corresponds to some extent with the pilot's perception of 'bad weather'. Heavy rainfall is also correlated with stormy, turbulent weather and thunderstorm activities, and is therefore able to capture many of the weather related hazards. The unit of measure of rainfall rate is mm per second.

At present, data used in CATS for this node has been collected only from Schiphol and therefore relates to take-off and landing from there. This probability distribution is not representative of worldwide averages, nor does it include the distribution of the rainfall rate that flight crew might "experience" during a flight. Strictly speaking, given a destination airport, management can not influence its weather condition. But management can influence whether there is unnecessary navigation through adverse weather that increases risk by means of:

- defining weather policy
- constant information exchange about weather
- enhancing communication between pilot, dispatcher, ATC
- commitment of the flight crew and enforcing use of the procedures to avoid the bad weather conditions en route
- training in pilot decision making for bad weather avoidance.

These aspects applied both in preflight and in-flight decision making.

Preflight Weather Planning

- An analysis that addresses weather considerations prior to operating over any new route or into any new airport (procedure)
- Collaborate with the ATC System Command Center for constant information exchange (communication and coordination)
- Provide weather information from authority approved source to the dispatcher and pilot (communication and coordination)
- Enhanced communication between pilot and dispatcher to maintain safe operational control (communication and coordination)
- Define minimum weather criteria to meet operational requirements and policies for preflight weather avoidance (e.g. alternate airport, choosing flight paths and landing routes) (procedure)
- Create a daily strategic plan of operations based on known or forecasted events two to six hours in the future (procedure)
- Ensure flight crew members, prior to each flight, to complete a review of weather information (including en-route and departure, destination and alternate airports) and to

En route Weather Avoidance

- Listen to ATIS and ASOS/AWOS broadcasts or ATC frequencies en route which help update and validate preflight weather information about conditions along route, when necessary, re-analyze flight plan (communication and coordination)
- Aircraft are equipped with an airborne weather radar system capable of detecting thunderstorms and other potentially hazardous weather conditions (man-machine interface)
- Ensure flight crew members, before operations in the proximity of adverse weather, openly discuss about weather condition, instructions, alternate airports, hazards and experience (communication and coordination)

Commitment

- Pilots supervise whether his/her colleague takes risk to go through the bad weather and take immediate action to correct deviations
- Management rewards strict adherence to weather-related procedures and takes disciplinary action against violations
- Management commitment to continuous improvement and joint training to develop collaborative solutions to weather constraint issues

Competente

- Train flight crew members to enhance their decision making in adverse weather and environmental conditions

Reference:http://www2.hf.faa.gov/weatherdecisionguide/in_flight.aspx

Man-machine interface- aircraft generation:

The quality of the interface between machine (the aircraft) and its human operator (the flight crew) has greatly improved over the years. This can be illustrated by comparing the cockpit of a first generation commercial jet transport aircraft like the De Havilland DH-106 Comet, with that of a modern jet airliner like the Boeing 777. When considering the effect of technological advances on safety of air transport it is common to consider four different generations of aircraft since the introduction of the jet engine. First generation aircraft are typically designed in the 1950s. Most of the aircraft were certified before 1965 according to British Civil Airworthiness Requirements (BCAR's) or other certification bases. Jet engines were still very new, and the aircraft had very limited cockpit automation, simple navigational aids and limited approach equipment.
Second generation aircraft, designed in the 1960s and 1970s, have more reliable engines. The aircraft were certified between 1965 and 1980, not yet based on common JAR-25/FAR-25 rules. Cockpit equipment is more advanced, with better auto pilots, auto throttles, flight directors and better navigational aids.
Third generation aircraft, designed in the 1980s and 1990s, typically show considerations for human factor aspects in the cockpit. Electronic Flight Instrument Systems (EFIS) and improved auto pilots are being used. Furthermore, the aircraft are equipped with ACMS data systems and high-by-pass engines designed according to higher certification standards.
Fourth generation aircraft like the Airbus A 320 and Boeing 777 have fully glass cockpits and digital fly-by-wire systems. Those different aircraft generations provide a convenient classification for the quality of the man-machine interface. Research has shown that the probability of flight crew error is significantly reducing for subsequent aircraft generations.

The only management influence directly on this operationalisation is the management policy towards purchase of new aircraft and phasing out old ones.

- Fleet composition

This influence does not require any expert judgement, since we already have estimates of how many aircraft of each generation there are and what effect the difference in generations has on error rate. To assess the effect of changing policy on fleet composition simply means changing the relative numbers of each generation.

Training- the number of days since the last type recurrent training

In training we would like to represent the quality, the content, as well as the amount of training that a flight crew member has received. Flight crew receive three types of training: technical knowledge, technical skills and non-technical skills. Technical knowledge is trained at ground school and technical skills are covered in the initial and recurrent type training. Training of non-technical skills is primarily covered in the Crew Resource Management (CRM) training. Both the content and the frequency of training are considered to be relevant factors. Because we are at this stage unable to quantify the quality of training in objective units, the number of days since the last type recurrent training will be used in the model to quantify 'training'[4].

|Management cannot do much about the number of days since last type recurrent training. They can only the influence the planned frequency of retraining, whether the crew stick to this frequency and the possibility of assigning two crew who have together at least a specified minimum of time since training – to avoid both having not had refresher training for a long time. We do not need expert judgement to assess the effect of these policies, since we already have data on the actual distribution of days since training and the estimates of the effect of this on error rate. The effect of implementation can simply be assessed by altering the actual distribution of days since training.

Competence

- Airline or authority changes the requirements of training period to hone the pilots' recency-of-experience

Competence-choose more experience & knowledgeable staff

- Airline screens the candidates for captain/first officer with an (enhanced) minimum level of technical competencies and skills, aviation experience, credentials and licenses, interpersonal skills

---

[4] With this operationalisation, technical knowledge, technical skills covered in the initial and CRM are not covered

Safety culture-not included in the model (it is left out of this exercise)

Safety culture is a complex, multidimensional issue that is not easily quantified. Hudson proposes the following five different stages of safety culture [Hudson 2003]:

1. Pathological: safety is a problem caused by workers. The main drivers are the business and a desire not to get caught by the regulator
2. Reactive: organisations start to take safety seriously but there is only action after incidents.
3. Calculative: Safety is driven by management systems, with much collection of data. Safety is still primarily driven by management and imposed rather than looked for by the workforce.
4. Proactive: with improved performance, the unexpected is a challenge. Workforce involvement starts to move the initiative away from a purely top down approach.
5. Generative: there is active participation at all levels. Safety is perceived to be an inherent part of the business. Organisations are characterised by chronic unease as a counter to complacency.

Safety culture is traditionally measured by means of questionnaires, see for instance [Von Thaden et al 2003] and [Kho et al 2005]. None of the methods for assessing safety culture meet the requirement that safety culture is expressed in objectively quantifiable units. Therefore safety culture is considered to be insufficiently made operational to be included in the current flight crew performance model. More research is needed on this topic and it is recommended to continue exploring possible ways to represent this aspect in future versions of the model.

References

Stein, E.S., Rosenberg, B.L. 1983. The measurement of pilot workload, Report No. DOT/FAA/EM-81/14, FAA Technical Center, Atlantic City Airport, New Jersey, USA.

Hart, S.G. 1987. The prediction and measurement of mental workload during space operations, paper presented at the NASA Space Life Sciences Symposium , Washington D.C., June 1987.

Lysaght, R.J., Hill, S.G., Dick, A.O., Plamondon, B.D., Linton, P.M., Wierwille, W.W., Zaklad, A.L., Bittner, A.C. & Wherry, R.J. 1989. Operator Workload: Comprehensive Review and Evaluation of Operator Workload Methodologies, Technical Report 851, United States Army Research Institute for the Behavioral and Social Sciences, USA.

Rosekind, M., Gander, P.H., Miller, D.L., Gregory, K.B., Smith, R.M., Weldon, K.J., Co, E.L., McNally, K.L., Lebacqz, J.V. 1994. Fatigue in operational settings: examples from the aviation environment, Human Factors, Vol 36, No. 2, p 327-338.

Simons, M., Valk, P.J.L. 1993. Review of human factors problems related to long distance and long endurance operation of aircraft. NATO-AGARD CP-547: Recent Advances in Long Range and Long Endurance Operation of Aircraft. Neuilly sur Seine: NATO-AGARD. p. 15/1-15/9.

Simons, M., Valk, P.J.L., de Ree, J.J.D., Veldhuijzen van Zanten, O.B.A. & D'Huyvetter, K. 1994. Quantity and quality of onboard and layover sleep: effects on crew performance and alertness. Report RD-31-94. Netherlands Aerospace Medical Centre, Soesterberg.

Simons, M., Valk, P.J.L. 1997. Effects of a Controlled Rest on the Flight Deck on Crew Performance and Alertness. Report: NLRGC 1997-B3. Netherlands Aerospace Medical Centre, Soesterberg.

Simons, M., Valk, P.J.L. 1998. Early starts: effects on sleep, alertness and vigilance. AGARD-CP-599; NATO-AGARD, Neuilly-sur-Seine, France. p. 6/1-6/5.

Valk, P.J.L., Simons, M. 1996. Effects of early reporting times and irregular work schedules on sleep, alertness, and performance of pilots engaged in short-haul operations. Report: NLRGC 1996-B2. Netherlands Aerospace medical Centre, Soesterberg.

JH Goode, 2003. Are pilots at risk of accidents due to fatigue? Journal of Safety Research, Vol. 34

VonThaden, T.L., Wiegmann, D.A., Mitchell, A.A., Sharma, G., Zhang, H. 2003. Safety culture in a regional airline, results from an aviation safety survey, 12th international symposium on Aviation Psychology, Dayton, Ohio, USA.

Hudson, P. 2003. Applying the lessons of high risk industries to health care, Quality and Safety in Health Care, 12, 7-12.

Kho, M.E., Carbone, J.M., Luicas, J., Cook, D.J. 2005. Safety Climate Survey: reliability of results from a multicenter ICU survey, in Quality and Safety in Health Care, 14, pp 273-288.

Caldwell, J. A. (1997). Fatigue in the aviation environment: An overview of the causes and effects as well as recommended countermeasures. Aviation, Space,&Environmental Medicine, 68, 932–938.

http://human-factors.arc.nasa.gov/zteam/
http://www.casa.gov.au/aoc/fatigue/regs/part02.htm
http://www.casa.gov.au/aoc/fatigue/regs/part02.htm
http://www.casa.gov.au/aoc/fatigue/regs/part02.htm
http://www.casa.gov.au/aoc/fatigue/regs/part02.htm
http://www2.hf.faa.gov/weatherdecisionguide/in_flight.aspx