The decision-making of commercial airline crews following an international pattern.

RENÉE M. PETRILLI, MATTHEW J.W. THOMAS, DREW DAWSON, GREGORY D. ROACH
Centre for Sleep Research, University of South Australia

Introduction
It is well established that crew decision-making is central to the safety and success of flight operations (Batt & O’Hare, 2005). Surprisingly, while there has been an increase in irregular and extended work hours in aviation, and a subsequent increase in the prevalence of fatigue (Caldwell, 2005), there has been limited progress in the understanding of the effects of fatigue on flight crew decision-making, particularly during critical decision events. As fatigue-related decision errors have been linked to several aviation incidents and accidents (e.g., NTSB, 1993; ATSB, 2006) it is of paramount importance to establish the impact of fatigue on the decision-making process of flight crew. Accordingly, the present study investigated the effects of prior duty history, and amount of actual sleep obtained on the decision-making performance of flight crew during a simulated critical decision event (CDE).

In general, previous research suggests that decision-making ability declines as fatigue levels increase (e.g., Harrison & Horne, 1999). Notably however, studies investigating pilot decision-making have typically used low-level cognitive performance tasks or low-fidelity flight simulators that may only be partially generalisable to the commercial airline environment. Moreover these studies have examined the behaviour of the individual and not the behaviour of the crew. Foushee et al. (1986) point out that the implications of fatigue may be quite different in an individual setting compared to a team environment such as a 2-pilot crew. Individual task environments omit factors such as crew co-ordination, leadership, and communication that may influence crew performance during safety-critical flight operations (Hawkins, 1987).

Only one study on fatigue and crew decision-making in a commercial airline environment was found in the literature (Foushee et al., 1986). Results of this study revealed that despite nominally fatigued crews reporting less sleep and higher levels of subjective fatigue associated with recent duty history compared to rested crews, somewhat paradoxically, these crews achieved better overall performance than low-fatigue crews. The authors surmised that the effects of familiarity between the Captain and the First Officer in the high-fatigue crews and/or strong operating procedures may have offset any fatigue-related performance decrements. If this was indeed the case, these results highlight the complex effects fatigue may have on decision-making performance in team environments.

The primary aim of the current study was to investigate the impact of fatigue associated with an international flying pattern on the decision-making of flight crew. A secondary aim was to identify potential strategies that may serve to protect flight crew performance from fatigue-related impairment. To guide the analysis and interpretation of results, we employed a prototypical model of naturalistic decision making - Klein’s (1993) recognition-primed decision (RPD) model.
Method

Participants
Participants were recruited from the Boeing 747-400 fleet of a commercial airline that operated international flights into and out of Australia. A total of 134 aircrew (67 Captains, 67 First Officers) participated in the study after bidding for a pattern that concluded with a study-related simulator session. All participants signed a consent form that signified their informed consent, confirmed their volunteer status, and stated that they understood their rights and obligations. All participants were assigned a unique identification code to ensure their anonymity. Participants were paid by the commercial airline at the standard rate for the simulator session that they attended in fulfillment of the experimental protocol. The study was approved by the University of South Australia Human Research Ethics Committee using guidelines established by the National Health and Medical Research Council of Australia.

Materials

Flight Simulator. Crew performance was assessed in a Boeing 747-400 full flight simulator (CAE Electronics Ltd), certified to the Civil Aviation Safety Authority (CASA) Level D prescribed in the FSD-1 (Operational Standards and Requirements-Approved Flight Simulators) Issue 4. At the time of the study, the commercial airline used these simulators for training and license renewals.

The Flight Scenario. Although several simulator scenarios were created, only one was used in the current study. Thus, every crew operated the same scenario. In the scenario crews were required to fly a full simulated flight from Sydney to Melbourne which has a flying time of approximately 60-70 minutes. The scenario included the normal pre-flight checks, take-off, cruise, descent, approach, and landing as well as various operational threats that the crew needed to manage. Results for crews’ management of the various operational threats are reported elsewhere. The current paper focuses explicitly on the critical decision event (CDE), which occurred at approximately 10 minutes prior to the top of descent. The CDE was designed to assess the important skill dimensions underlying decision-making. There was no single (in) correct resolution to the CDE, which is described below.

At the outset, crews plan to land on Runway 16 in Melbourne. The aircraft is dispatched with the Engine No. 3 Thrust Reverser “locked out” which means the plane has reduced braking capability on the runway. Crews receive a valid terminal area forecast (TAF) for Melbourne that indicates that the visibility and cloudbase are acceptable and above the alternate criteria for landing.

- At approximately 10 minutes prior to the top of descent, Air Traffic Control (ATC) issues a revised Aerodrome Terminal Information Service (ATIS) (i.e., change of weather conditions) for Melbourne, informing the crew that the weather conditions have changed in Melbourne.
- The new ATIS indicates that winds in Melbourne have increased, with the implication they are now above crosswind limitations for Runway 16 such that landing on this runway is no longer legal.

The change in weather presents the crew with three critical issues that they must resolve in order to decide whether to (1) continue to, or (2) divert from the destination airport Melbourne:
1. Can they land on the original runway in Melbourne (RWY 16) given that there is now a strong crosswind and the runway is wet?
2. Can they land on the perpendicular runway in Melbourne (RWY 27) which has no crosswind, but it is a short runway, it is wet, and the plane has reduced braking capability due to the Engine No. 3 Thrust Reverser being locked out?
3. Can they land at a different airport considering the weather at other airports and fuel constraints?

Dependent variables

Decision-Making. To assess crew decision-making performance during the CDE, an observer viewed and coded each simulator session from the video recordings. Several variables extracted from a list of measures of crew performance (from the video analyses) were chosen for the analyses. It was
proposed that the chosen variables had the highest relevancy to crew decision-making that readily tap the following areas of crew decision-making based on Klein’s (1993) recognition-primed decision (RPD) model: (i) situation awareness; (ii) options and planning; (iii) decision implementation; and (iv) evaluation. The decision-making variables are presented within the framework of the RPD (Table 1).

<table>
<thead>
<tr>
<th>Decision-Making Stage</th>
<th>Dependent Variables</th>
</tr>
</thead>
</table>
| Situation Awareness    | Cross-check Figures? (categorical) - Did crews cross-checked the landing figures?  
|                       | Obtain Melbourne Weathers/Trend Forecasts (categorical) – Did crews request weather information including forecasts and trends?  
|                       | Determine Melbourne below alternate criteria (categorical) - Did crews determine that Melbourne was below alternate criteria such that an alternate airport was necessary? |
| Options and Planning   | Number of Options (continuous) – Number of options considered by crews.  
|                       | Determine Runway 16 Available (categorical) - Did crews determine that Runway 16 would be available if the wind dropped when heading for Runway 27? |
| Decision Implementation| Time to Finalise Decision (continuous) – Time taken for crews to verbally decide their plan of action.  
|                       | Time to Positive Action (continuous) - Time taken for crews to take positive action towards the decision (e.g., execution of FMC or request diversion).  
|                       | Divert from MEL? (categorical) - Did crews divert to another airport instead of continuing to Melbourne? |
| Decision Evaluation    | Review decision? (categorical) – Did the crews review the decision? |

Independent Variables

Sleep/Wake Schedules. Pilots’ sleep/wake schedules were obtained using sleep diaries and activity monitors. Pilots kept self-recorded sleep diaries for the entire study period (i.e., approximately 15 days) where they recorded information for each sleep period (including all in-flight sleep periods) including self-rating their level of fatigue using the Samn-Perelli Fatigue Checklist (Samn & Perelli, 1982), and perceived sleep quality using a 5-point Likert Scale.

Objective sleep/wake schedules were assessed using activity monitors (Mini MitterTM, Sunriver, Oregon), which were also worn by each pilot for the entire study period. Activity monitors are devices worn like a wristwatch on the wrist that allow for 24-hour recordings of activity (see Ancoli-Israel et al., 2003).

The independent variable derived from the sleep diaries and activity monitors was Sleep in prior 24h (in hours) - the amount of sleep that a participant had obtained in the 24 hours prior to the start of the simulator session.

Work/Rest Schedules. Pilots’ work/rest schedules were obtained from duty diaries that were kept for the entire study period (i.e., approximately 15 days). In the duty diaries, participants recorded information for every work period, which included the on-blocks and off-blocks time (i.e., start and end) of each flight, the origin and destination ports, and subjective fatigue using the Samn-Perelli Fatigue Checklist before and after each flight. Time was recorded as universal time, coordinated (UTC). The independent variable derived from the duty diary was duty – whether the crews were rested or non-rested (see below).

Psychomotor Vigilance Task. Pilots’ response times were measured using a PalmPilot version of the psychomotor vigilance task (PVT, Ambulatory Monitoring Inc.), developed by the Walter Reed Army Institute of Research (see Thorne et al., 2005). This version of the task has been validated in several studies (e.g., Lamond et al., 2005). The independent measure derived from the PVT was
mean response speed, expressed as the mean reciprocal response time multiplied by 1000, as per standard methodology. Lower scores indicated a greater level of impairment.

**Subjective Fatigue.** Subjective fatigue was assessed using the Samn-Perelli Fatigue Checklist (Samn & Perelli, 1982). The Samn-Perelli is a 7-point Likert scale with: 1 = “Fully alert, wide awake”; 2 = “Very lively, responsive, but not at peak”; 3 = “Okay, somewhat fresh”; 4 = “A little tired, less than fresh”; 5 = “Moderately tired, let down”; 6 = “Extremely tired, very difficult to concentrate”; and 7 = “Completely exhausted, unable to function effectively”. The independent measure derived from the Samn-Perelli Fatigue Checklist was self-rated fatigue. Higher scores indicated a higher level of subjective fatigue.

**Design and Procedure**
The study employed a completely randomised design. Sixty-seven crews (Captain and First Officer) operated a B747-400 simulator either (i) after having at least four consecutive days free of duty after an international pattern (rested), or (ii) immediately after the final landing in Sydney at the end of an international pattern (non-rested). Ultimately, there were:

- 21 rested crews
- 22 non-rested crews that had returned from patterns to Europe
- 21 non-rested crews that had returned from patterns to the United States
- 3 non-rested crews that had returned from patterns to South Africa

There were no significant differences between the rested and non-rested crews in terms of the total number of hours the pilots had flown B747 aircraft and the total number of flying hours. To minimise the impact of crew familiarity effects on crew performance, we ensured that the rested and non-rested crews’ Captains and First Officers had flown together during their last international pattern.

The simulator sessions were identical for the rested and non-rested crews (see Figure 1).

- **Pre-flight Testing.** At the start of the simulator session, participants completed the pre-simulator questionnaires and a 5-minute PVT (approximately 15 minutes).
- **Pre-flight Planning.** Participants began their pre-flight planning (i.e. fuel, route, weather, etc.) approximately 30-minutes prior to entering the simulator.
- **Simulator Session.** The simulator session involved a single flight sector (Sydney-Melbourne) with all stages of flight – pre-flight, take-off and climb, cruise, descent and approach, and landing (approximately 2 hours). In all cases, the Captain was the pilot flying.
- **Post-flight Testing.** At the end of the simulator session, participants completed the post-simulator questionnaires and a 5-minute PVT (approximately 10 minutes).

Post-flight interviews were conducted with each crew that lasted approximately 30 minutes. Interviews were tape recorded with the permission of the crew and subsequently transcribed.

**Analyses**
To assess the impact of fatigue on the decision-making of flight crew, independent variables corresponded to one of the first three levels of the fatigue hazard trajectory (Dawson & McCulloch, 2005).

- **Level 1:** Recent Duty History - prior sleep opportunity based on recent duty history.
- **Level 2:** Actual sleep – amount of sleep obtained in the 24 hours prior to the start of the simulator session.
- Level 3: Self-rated Fatigue – self-rated fatigue using the Samn-Perelli Fatigue Checklist prior to the start of the simulator session.
- Level 3: PVT Response Speed – inverse reaction time (relative to baseline) on the Psychomotor Vigilance Task prior to the start of the simulator session.

For the analyses, crews were allocated to groups (low-, moderate-, and high-fatigue) on the basis of the independent variables. Unpaired samples t-tests and chi-squared cross-tabulations were used to examine the impact of the independent variables on crew performance. Themes extracted from the interview analyses were coded based on a grounded theory approach employed by Chrichton and Flin (2004). Two raters not part of the simulator study, but expert in the areas of fatigue and performance, each cross-coded a sample of 15 randomly selected transcripts. An average Cohen’s Kappa of $K = 0.82$ indicated high inter-rater reliability.

**RESULTS**

The results relating to crew decision-making are organised based on the four stages of the decision-making process: (i) situation awareness, (ii) options and planning, (iii) decision implementation, and (iv) evaluation.

**Situation Awareness.** In all analyses moderate- and high-fatigue crews indicated better performance compared to low-fatigue crews, suggesting that fatigued flight crew may have employed a high-effort strategy to cope with fatigue (i.e., they were more thorough/meticulous). Analyses indicated that whether crews cross-checked the landing figures (Cross-check Figures?) was significantly affected by recent duty ($\chi^2[1]=6.2, p<.05$), and the amount of sleep in the prior 24 hours of the Captain ($\chi^2[2]=7.0, p<.05$), and the self-rated fatigue of the both the Captain and First Officer ($\chi^2[2]=6.7, p<.05$). Whether crews obtained information regarding Melbourne weathers or trend forecasts (Obtain Mel Weathers/TTF?) was significantly affected by recent duty ($\chi^2[2]=9.8, p<.01$). Notably however, whether crews determined Melbourne was below alternate criteria (i.e., it was necessary to establish an alternate airport before landing in Melbourne- Determine MEL below alternate?) was not affected by fatigue.

**Options and Planning.** The options and planning stage of the decision-making process remained unaffected by the Level 1, Level 2, or Level 3 fatigue indicators. Specifically, the analyses indicated no significant differences between groups in terms of the number of options considered (Number of Options) or whether crews determined that Runway 16 would be available if the wind dropped (Determine RWY16 Available?).

**Decision Implementation.** The analyses revealed that the implementation of the decision was highly sensitive the effects of recent duty history ($t_{33}=2.0, p<.05$), and the self-rated fatigue of the Captain ($t_{24}=2.1, p<.05$). Accordingly, fatigued crews tended to take longer to finalise their decision (Time to Finalise Decision). Figure 2 illustrates the time to finalise the decision with respect to the self-rated fatigue of the Captain. The approximate time at which crews should commence the descent-approach-landing phase of flight is included in the figure to illustrate the operational significance of taking longer to resolve the critical decision event.

![Figure 2](image_url)

Figure 2 represents the time taken for crews to finalise the critical decision event when allocated to groups on the basis of the self-rated fatigue of the Captain (±SD). Asterisk indicates significant difference from low-fatigue crews ($p<.05$)
Furthermore, with respect to crew diversion (Divert from Melbourne), a greater proportion of non-rested crews chose to divert from the destination airport, Melbourne compared to rested crews ($\chi^2[2]=6.7, p<.05$). This may indicate that fatigued crews may tend to opt for more conservative options during complex decision-making. However, the time to take positive action towards the decision (Time to Positive Action) was not significantly affected by fatigue.

**Evaluation.** The analyses indicated that whether crews evaluated the decision (Review Decision?) was not significantly affected by the Level 1, Level 2, or Level 3 fatigue indicators.

**Performance Protection Strategies.** All pilots confirmed they employed performance protection strategies during normal flight operations to cope with fatigue. When asked whether performance protection strategies were used during the simulator session, chi-square cross-tabulations indicated that 82.3% of non-rested crews stated they had employed these strategies, whereas only 21.1% of rested crews stated they had employed these strategies ($\chi^2[1] = 19.51, p<.001$). Several performance protection strategies were elicited from the interviews. These have been classified into three categories (i) internal- dependence on self to maintain performance, (ii) external- dependence on other (machine or human) to maintain performance, (iii) general- countermeasures employed to counteract fatigue. Table 2 presents these findings.

**Table 2. Flight Crew Fatigue Protection Strategies During Normal Flight Operations.**

<table>
<thead>
<tr>
<th>Performance Protection Strategy</th>
<th>Transcript Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal</td>
<td></td>
</tr>
<tr>
<td>Reliance on SOPs</td>
<td>“You use standard ops to get through.”</td>
</tr>
<tr>
<td>Lists and mnemonics</td>
<td>“I write a note to myself and stick it under the control.”</td>
</tr>
<tr>
<td>Self-pacing</td>
<td>“[I] slow it down, concentrate, and just do the essentials…[I] don't try to do too many things.”</td>
</tr>
<tr>
<td>Expectation of error</td>
<td>“I tend to be more, well, particularly methodological and trust nothing.”</td>
</tr>
<tr>
<td>Increased cross-checking</td>
<td>“I check things three or four times over.”</td>
</tr>
<tr>
<td>Increased pre-planning</td>
<td>“I tend to prepare well in advance when I'm tired.”</td>
</tr>
<tr>
<td>Dependence on adrenaline</td>
<td>“At the time of descent, adrenaline seems to kick in.”</td>
</tr>
<tr>
<td></td>
<td>“I just try and crank myself up.”</td>
</tr>
<tr>
<td>Increase workload</td>
<td>“I just make myself busy.”</td>
</tr>
<tr>
<td>Fatigue acknowledgment</td>
<td>“[I] tell myself that I'm tired and try to be a bit more alert.”</td>
</tr>
<tr>
<td>External</td>
<td></td>
</tr>
<tr>
<td>Support from crewmembers</td>
<td>“…getting other people to check or evaluate, ensure that what you are thinking is right and getting some feedback from other people to make sure you are on the right track.”</td>
</tr>
<tr>
<td>Dependence on automatics</td>
<td>“[I’m] more inclined to use the automatics. To hand fly an aircraft takes about 85% of your brain power and if you can take that away then you've got time to sit back and absorb what's happening more easily.”</td>
</tr>
<tr>
<td>General</td>
<td></td>
</tr>
<tr>
<td>Anticipation of fatigue</td>
<td>“[I] try and prepare, anticipate fatigue so as to avoid the rush and to minimise pressure at later stages of the flight.”</td>
</tr>
<tr>
<td>Caffeine consumption</td>
<td>“[I’ll have]…probably 3 or 4 [coffees] on a sector, and then drink water and even have a Coke with a meal or something like that, a bit of caffeine and sugar.”</td>
</tr>
<tr>
<td>Caffeine avoidance</td>
<td>“[I] avoid coffee, makes me feel worse… you find that your brain's going so fast, but its not going in the right direction all the time and you're actually not making good decisions.”</td>
</tr>
<tr>
<td>Napping</td>
<td>“Its amazing that if you are in that extremely fatigued state, its amazing what a 10 minute nap can do.”</td>
</tr>
<tr>
<td>Food/Sweets/Chocolate</td>
<td>“I’ll often eat when I’m not necessarily hungry but when I’m starting to feel tired, it just gives you something to do for a while…extra calories and some carbs to get the body working a little bit.”</td>
</tr>
<tr>
<td>Physical activities</td>
<td>“[I] get up away from the chair and stretch my legs, splash some water on my face”.</td>
</tr>
<tr>
<td></td>
<td>“Exercise.”</td>
</tr>
</tbody>
</table>

**DISCUSSION**

In focussing on the mechanisms underpinning crew performance during a critical event with no single correct solution, this study has taken a different approach to any other investigating fatigue and flight crew decision-making in naturalistic environments. Variables chosen for the analyses were extracted from a list of several measures of crew performance. It was proposed that the chosen variables had
the highest relevancy to crew decision-making and readily tap the following areas of the decision-making process: (i) situation awareness; (ii) option assessment; (iii) decision implementation; and (iv) evaluation. Viewing the variables within a general naturalistic decision-making framework helped to direct our focus and facilitated the interpretation of the impact of fatigue on the various stages of crew decision-making. Further to this, it provided a general platform from which results from the current study may best be generalised to other high-risk environments.

With respect to situation awareness, results of the present study indicate that some aspects of situation awareness are sensitive to fatigue, particularly those aspects related to the acquisition and maintenance of situation awareness. More than this though, the results indicate that these aspects in fact improve under conditions of fatigue. For example, when crews were allocated to groups on the basis of Level 1 (prior duty), Level 2 (prior sleep), and Level 3 (subjective fatigue, PVT performance) fatigue indicators, a greater proportion of moderate- and high-fatigue crews cross-checked the landing figures compared to low-fatigue crews. Similarly, when crews were allocated to groups on the basis of Level 3 fatigue indicators, a greater proportion of moderate- and high-fatigue crews acquired information regarding the weather and trend forecasts for the destination airport Melbourne compared to low-fatigue crews. From a safety perspective, calculating the landing figures during the scenario enabled crews to determine whether the length of the desired runway (27) was sufficiently long enough to legally land the aircraft given weather conditions and aircraft constraints.

Results from the current study also indicated that on the basis of Level 1, Level 2, and Level 3 fatigue indicators, a greater proportion of moderate-fatigue and high-fatigue crews diverted from Melbourne compared to low-fatigue crews. Compared to returning to the origin airport Sydney, or diverting to another airport, the destination airport Melbourne contained the greatest level of uncertainty and risk as it involved adverse weather conditions and landing on a limited runway with a reduced braking capability. Several studies have found that decision-makers have a high aversion to risk when there is an element of uncertainty (Hendrickx & Vlek 1991; Orasanu, Fisher, and Davidson, 2004). The current findings add to this notion further by suggesting that fatigued flight crew may be more likely to avoid risky options and divert. Indeed conservative decision-making was identified by crewmembers as a primary consequence of fatigue during the post-flight interviews. Specifically, crewmembers indicated that during the simulated flight they preferred to stay within their “comfort zone” and divert rather then continue to a risky destination.

Another important finding with respect to the implementation of the decision was that the time taken to finalise the decision was particularly sensitive to fatigue at Levels 1, Level 2, and Level 3 of the fatigue hazard trajectory. Our results indicated that moderate- and high-fatigue crews took 34-42% longer to finalise the decision compared to low-fatigue crews. From an operational perspective, taking longer to make decisions may compromise flight safety as it can lead to greater time pressures at later stages of flight, particularly during the high-workload stages of descent, approach, and landing. Indeed, Urban et al. (1996) found that under high-time pressure teams perform significantly worse and poorer under high workload demands. Furthermore, as the critical decision event in the current study did not need to be resolved within a pre-determined time-frame, the effects of fatigue on flight crew decision-making may be exacerbated under conditions of time stress. Clearly, this is an important area of future investigation and suggests that crews should start preparing for the descent, approach, and landing phases of flight, earlier than 10 minutes prior to the top of descent to avoid situations associated with time pressures. Furthermore, while we attempted to control for the effects of familiarity on crew performance by ensuring that both rested and non-rested crews’ Captains and First Officers had flown together during their last international pattern, this factor may still have influenced crew performance. Consequently, future research should attempt to investigate interaction effects between crew familiarity and fatigue on crew decision-making.

Human factors research has focused extensively on the decision-making of individuals and teams, and the consequences of ineffective decisions are well understood in the aviation environment (ATSB, 2006; NTSB, 1993). This study has provided new knowledge regarding the processes underlying flight crew decision-making and has furthered our understanding of the impact of fatigue on the decision-making process. Specifically, the findings demonstrate that prior sleep opportunity associated with recent duty history (Level 1 fatigue control), actual prior sleep obtained (Level 2 fatigue control), subjective fatigue, and response times of crewmembers (Level 3 fatigue controls),
can influence the crew decision-making process. In particular, flight crew tend to take longer to make decisions, are more conservative with their decision-making, and appear to apply performance protection strategies to cope with fatigue. Our interviews with crewmembers also highlighted that the impact of fatigue on operational flight performance is an important issue for crewmembers, and that a range of performance protection strategies are commonly employed to deal with the effects of fatigue. In light of these new findings, and given the importance of sleep and its well-established relation to subjective and objective fatigue (e.g., Lamond et al., 2005), it is clear that ways to improve the amount and quality of sleep before and during duty periods need to be considered. Plainly this would help crews to make effective decisions under a wider range of conditions particularly when encountering unexpected critical events.

References