CARE Innovative Action

Cognitive Streaming Project

Report on Work Package 4: 'Real-World' Effects

Ref: CARE-IA-CS-NLR-WP4-D4-02-1.0.doc

Piet Hoogeboom, Eamonn Hanson, Martin Joosse, Helen Hodgetts, Dylan Jones, Alan Salmoni, Eric Farmer and Sonja Straussberger

DOCUMENT REVIEW

Version	Date	Description	Modifications
A	15 April 2004	Incomplete draft for comment at progress meeting	
B	4 June 2004	Second draft for comment	Changes/additions as described in minutes of meeting PM5
C	2 July 2004	Third draft for comment	Changes/additions as suggested by Mick van Gool
1.0	22 July 2004	Final version	Minor changes as suggested by Mick van Gool

DISTRIBUTION LIST

Consortium		EUROCONTROL	
Eric Farmer	QinetiQ	Mick van Gool	EUROCONTROL CARE Manager
Dylan Jones	Cardiff	Dirk Schaefer	EUROCONTROL ERC
Helen Hodgetts	Cardiff	Oliver Straeter	EUROCONTROL DAS/HUM
Alan Salmoni	Cardiff	Sven Fraenkel	EUROCONTROL INTEGRA Action Manager
Fabrice Parmentier	Cardiff	Marco Brusati	EC DG Research
Herman Nijhuis	NLR	Tjien-Khoen Liem	EC DG Research
Martin Joosse	NLR	Poul Erik Jorgensen	EUROCONTROL DAS/ATS
Harrie Bohnen	NLR	Sonja Straussberger	EUROCONTROL ERC
Piet Hoogeboom	NLR		
Victoria Chapman	QinetiQ		
Adam Brownson	QinetiQ		

Table of Contents

GLOSSARY

AMTE	Aviation Multi-Tasking Environment
ANOVA	Analysis of Variance
ATC	Air Traffic Control
CARE	Co-operative Actions of R&D in EUROCONTROL
df	Degrees of Freedom (in statistical tests)
DME	Distance Measuring Equipment
ECG	Electrocardiogram
EDA	Electro-Dermal Activity
EEG	Electroencephalogram
EFIS	Electronic Flight Information System
EFRP	Eye Fixation Related Potential
EL	Experiment leader
EM	Experiment Manager
EOG	Electro-oculogram
FL	Flight Level
GPWS	Ground Proximity Warning System
GRACE	Generic Research Aircraft Cockpit Environment
GSR	Galvanic Skin Response
HDG	Heading
HEART	Human factors Evaluations, data Analysis and Reduction Techniques
ISE	Irrelevant Speech Effect
LSD	Least Significant Difference
MFF	Mediterranean Free Flight
MSE	Mean Square Error
NASA	National Aeronautics & Space Administration
NDB	Non-Directional Beacon
NLR	Nationaal Lucht en-Ruimtevaartlaboratorium (National Aerospace Laboratory)
NOTAMS	Notices to Airmen
p	probability (in statistical tests, the probability that a given effect occurred by chance; by convention, p values of less than 0.05 are considered statistically significant)
PF	Pilot Flying
PNF	Pilot Not Flying
RAM	Random Access Memory
REDFA	Entry Point of the Amsterdam Flight Information Region
RSME	Rating Scale Mental Effort
RT	Reaction Time
R/T	Radio/Telephony
RWY	Runway
SA	Situation(al) Awareness
SCL	Skin Conductance Level
SD	Standard Deviation
SPL	Schiphol
SPO_2	Arterial oxyhaemoglobin saturation
STAR	Standard Terminal Arrival Route
SUGOL	An Initial Approach Fix for Schiphol Airport
SULUT	A Waypoint on the Route into Schiphol
TLX	Task Load indeX
VHF	Very High Frequency
VMC	Visual Meteorological Conditions
VOR	VHF Omnidirectional Range
WP	Work Package

LIST OF FIGURES AND TABLES

ABSTRACT

This EUROCONTROL CARE project aims to demonstrate how a new theoretical approach to understanding human information processing — 'cognitive streaming' — can be applied to problems facing aircrew and air traffic controllers. Experiments in Work Package (WP) 2 had shown that irrelevant sound disrupted many aviation-relevant tasks in the laboratory — only a psychomotor task (tracking) was unaffected. The aim of WP4 was to determine the effects of the party line in realistic simulated flight, using the NLR GRACE facility.

Effort and mental workload ratings confirmed that the simulator task imposed a relatively high workload. Crews judged that the party line produced a performance impairment of about 25%, and increased the level of distraction, but all flights were perceived as being unaffected regarding safety. Although largely irrelevant to the task at hand, the additional stream of background information represented by the party line increased ratings of workload, mental activity, mental effort and time pressure. The party line condition created a multitasking situation in which more effort was required to co-ordinate two behavioural streams; the effect was greater for the pilot not flying (PNF), responsible for management of R/T communication, than for the pilot flying (PF). The level of distraction from R/T communications was considered to be greater for the PNF.

A variety of physiological measures were recorded. These data indicated that workload was higher for the PF than for the PNF. Heart rate increased steadily for the PF throughout the duration of the flight, but there was little effect for the PNF. Heart rate variability and blink rate were lower for PF than for PNF, supporting the notion of higher workload. The party line had some effect on physiological response. For example, skin conductance was slightly higher — indicating higher physiological activation — in the presence of party line. $SP0_2$ (arterial oxyhaemoglobin saturation) was significantly higher in the PF party line condition (suggesting lower workload). However, the physiological effects were much less marked than those evident in the subjective data, and in some cases even seem to contradict the Cognitive Streaming Theory.

The simulator-based performance results showed statistically significant effects, which were mostly related to the flight phase. In line with the results of WP2, the party line did not affect the flight path tracking accuracy as determined from movements in aircraft pitch and roll angles. Also some contrasting results were found. For instance, the distance between runway threshold and aircraft touchdown point showed lower deviation when the party line was absent. Differences between Captains and First Officers were also observed, but again the effects were inconsistent.

Observational data, including timings, provided a further objective basis for examination of the effects of the party line. The time required for the descent checklist was increased by the party line. Moreover, slightly more calls were missed, queried or incorrectly read back in the party line than the no party line condition, although there were insufficient data of this kind to perform statistical analyses. The in-flight observer noted subjective observations of distraction, such as the PNF visibly trying to attend to both the PF and the background R/T. Omissions and repetitions of checklist items seemed more dependent upon *interruption* of the checklist than upon mere *distraction* by background R/T. The recent literature reviewed in the WP3 report supports the notion that interruption is a flight deck issue that merits much more attention.

Differences in subjective opinion according to pilot rank were noted. For Captains, the party line increased many of the workload-related ratings, but these effects were absent for First Officers. Captains were affected by the background R/T in both PF and PNF roles, whereas the First Officers generally reported that the party line had an effect only when they were acting as PNF. During debriefing, the Captains claimed to monitor the party line in both flying roles, but First Officers attended to the R/T only as PNF. First Officers were also more distracted by the party line than Captains, and were more likely to report that their performance was impaired by background R/T. These differences may reflect flight experience and/or overall responsibility for the proper execution of the flight.

This trial represents the first substantial 'real-world' test of the cognitive streaming model. The implications for further development of the model are discussed.

General Introduction

1. INTRODUCTION

1.1. Background

A pilot study was conducted as part of the Eurocontrol CARE programme (Jones & Farmer, 2001) to demonstrate how a new theoretical approach to understanding human information processing — 'cognitive streaming' — could be applied to problems facing aircrew and air traffic controllers as traffic density increases. Cognitive streaming (e.g., Jones et al, 1997; Macken et al, 2003) is a theoretical framework for human information processing when auditory and visual information is combined. It was developed to understand the adverse effects of irrelevant background sound on a variety of tasks, particularly those that involve the retention of order in short-term memory. This approach postulates that interference between tasks occurs when they draw upon the same mental process (such as keeping track of order). It therefore differs from multiple resource theory (Wickens, 1992), which assumes that specialised 'resources', each with its own capacity limitation, can be distinguished.

1.2. Structure of the Cognitive Streaming project

The study described here is being conducted as part of the EUROCONTROL CARE Innovative Cognitive Streaming project by QinetiQ, Cardiff University, and NLR. This report describes Work Package (WP) 4 of the main study on cognitive streaming that follows on from the preliminary work summarised above. The Cognitive Streaming project is divided into four WPs.

WP1: Possible usefulness of irrelevant sound. The results of the pilot study confirmed that irrelevant sound will be registered, and have a damaging effect on task performance. The objective of the studies in this work package was to determine whether the information registered is useful to the individual (Jones, Parmentier, Houghton, Pope, & Farmer, 2002).

WP2: Range of tasks affected by irrelevant sound. The objective was to determine the extent to which the effects found in the pilot study generalise to other tasks (Pope, Houghton, Jones, & Parmentier, 2003).

WP3: Prevention of performance impairment. The objective of this WP was to determine how the disruptive effects of irrelevant sound could be avoided. Task analyses were performed upon flights, both from videotape and from direct flight deck observation, and a wide range of sources in the literature was consulted (Farmer, Chapman, Brownson, Thompson, & Jones, 2003).

WP4: Real-world effects. The objective of this WP, reported here, was to examine the extent to which the effects observed in the earlier work packages could be demonstrated in realistic aviation tasks. Most previous work on the Irrelevant Sound Effect (ISE) has been based upon simple laboratory tasks. WP4 provided an opportunity to assess the disruptive influence of irrelevant sound in an applied setting.

1.3. Aims of Work Package 4

The cognitive streaming model, and the results of WP2 in particular, suggests that the party line should interfere with many flight activities. Results using the Aviation Multi-Tasking Environment (AMTE) battery were as follows:

- Conflict detection: a general trend toward higher hit rates and lower false alarm rates in the quiet condition than in the two irrelevant sound conditions; correct response times were also longer with irrelevant sound
- Communication: accuracy was lower in both irrelevant sound conditions than in the quiet condition
- Visual monitoring: irrelevant sound increased reaction times to the monitoring events, particularly in the irrelevant speech condition
- Tracking: no significant effects of irrelevant sound on tracking performance

The overall aim of the study conducted in WP4 was to examine the extent to which effects observed in the laboratory could be demonstrated when flight crews were subjected to realistic flying tasks in a high-fidelity flight simulator.

1.4. Trial set-up

The NLR Generic Research Aircraft Cockpit Environment (GRACE) was selected as the environment for the study. As stated earlier, the aim of the trial was to examine the effect of the party line under realistic flying conditions. The specific objectives selected were:

- to expose crews to two conditions: 'minimal-required' Radio Telecommunication (R/T) as a control condition, and this R/T plus the party line as the experimental condition
- to collect performance data from the simulator during final descent, approach and landing at Schiphol Amsterdam airport
- to collect subjective data on performance, workload and reactions to the experiment from the PF and the PNF, by means of questionnaires and structured interviews
- to collect physiological data such as heart rate variability, blink rate, arterial oxyhaemoglobin saturation (SPO_2), galvanic skin response (GSR) and electroencephalography (EEG)

A draft test plan was produced by NLR, and a 'dry run' was held in the GRACE simulator in December 2003. The dry run led to many enhancements to the protocol, including:

- Type of task: a manual non-precision approach task, corresponding to a medium-to-high workload task.
- Choice of the appropriate questionnaires
- Identification of critical events during the simulated flight
- Performance measures to be derived from the simulation
- The session duration and number of simulation runs
- The work-share between partners

1.5. Structure of the report

The method, including facilities and material, the experimental design, and the procedure, is described in Section 2. Section 3 presents the results: flight data, questionnaire responses, physiological measures, and observational data. Section 4 comprises discussion of the results, including their implications for cognitive streaming theory. Following the References (Section 5), a series of appendices provides some detailed information on the trial, including results of formal statistical analyses.

2. METHOD

2.1. Facilities and Materials

2.1.1. The Generic Research Aircraft Cockpit Environment

GRACE (Figure 1) is a generic fixed-base flight simulator, which can be configured as a modern large airliner with or without side-stick control, and with two or four engines. In the fall of 2004, motion will be added to GRACE, making it a full research flight simulator. All software models (aircraft and systems) driving the simulator are developed and maintained by NLR.

For the Cognitive Streaming project, the simulator was configured as a Boeing 747, with Boeing 747-400 enhanced EFIS displays and a simulated Boeing 747-400 Flight Management System. Pilot positions and system controls were fully representative for the simulated aircraft type, including appropriate force feedback on the control yoke by the control loading system.

Figure 1: Generic Research Aircraft Cockpit Environment (GRACE)

During the trials records were kept of all relevant aircraft parameters such as altitude, position and system status. No malfunctions were simulated. Since manual approaches were performed, weather conditions were such that the runway was clearly visible in the computer-generated outside view. Pilots were instructed to disconnect the autopilot, but were allowed to use the flight director to help them to maintain the appropriate heading and altitude/vertical speed.

2.1.2. Questionnaires

The subjective evaluation was based upon the following questionnaires:

- Rating Scale Mental Effort (RSME) with verbal labels; see Appendix B.
- Rating scales based upon some of the NASA Task Load indeX (TLX) scales, with additional items; see Appendix B.

Before the experiment started, the pilots filled in a set of questionnaires. The information was used to provide biographical data and to assist in the analysis of physiological data:

- Experience and Biographical data
- Use of stimulants and medication
- Length of sleep

2.1.3. Video

For verification and back-up purposes, audio recordings were made of crew communication in the cockpit and the radiotelephony (RT, both with ATC and background party-line sounds). The audio recording was integrated with a low-fidelity video recording. The location of the video camera provided a diagonal front view of the crewmembers.

2.1.4. Physiological

Physiological signals were measured with a VitaPort-3 system. This portable data recorder is capable of registering several external analogue signals such as EEG at varying sampling frequencies. Each signal is read through a separate channel, pre-processed and stored on a RAM card. A similar VitaPort-2 system is routinely used in desktop, ATC, and moving-base flight simulations at NLR.

The following physiological signals were collected:

- ECG for the derivation of heart rate and heart rate variability
- Respiration (one cable)
- Vertical and horizontal electro-oculogram (EOG)
- Synchronisation bit from GRACE
- SPO_2 (arterial oxyhaemoglobin saturation, captain only)
- GSR (Galvanic Skin Response, First Officer only)
- EEG (Electroencephalogram, locations: Cz, Pz and Oz)
- Controller voice (via VOX) and controller push-button (VOX was used to synchronise controller voice with the EEG signal)

Due to the amount and type of recorded data, it was necessary to use separate data recorders. This requirement necessitates off-line synchronisation of the physiological signals, video and controller communication registrations with the data from the simulator. To achieve this, GRACE sent status information including a unique synchronisation bit to all systems.

2.2. Scenario and Flight Measures

2.2.1. General description of flight to Amsterdam Airport Schiphol

To increase the probability of demonstrating performance effects of the party line, a relatively high incidence of R/T with other aircraft was presented. The traffic simulated was visible on the NAV display and in the outside view. During the pilot study, it became apparent that the runs were too predictable, workload subjectively decreasing with familiarity. Increased variability between runs was therefore introduced in factors such as call signs of traffic, wind conditions and ATC directives. In the experiment, three almost similar scenarios were used, with different call signs and wind conditions.

The experimental run commenced at FL140 when the aircraft was flying level, at the SULUT waypoint on the REDFA Standard Instrument Arrival Route (STAR) into Schiphol (Figure 1). Distance to the runway was approximately 50nm. The final approach and landing were performed manually. Crews were asked to perform a non-precision approach, which was representative of a medium-to-high workload tracking task. Each approach and landing started with the completion of the descent checklist. The total segment lasted about 10–15 minutes.
The scenarios represented a routine flight arriving at Amsterdam Airport Schiphol according the following flight plan (see Figures 2 & 3):

INITIAL POSITION: FL140, 300kias, at SULUT
REDFA ARRIVAL
    SUGOL
    30nm SPL
    15nm SPL
RADAR VECTORS to INTERCEPT
    NDB DME approach RWY 18C

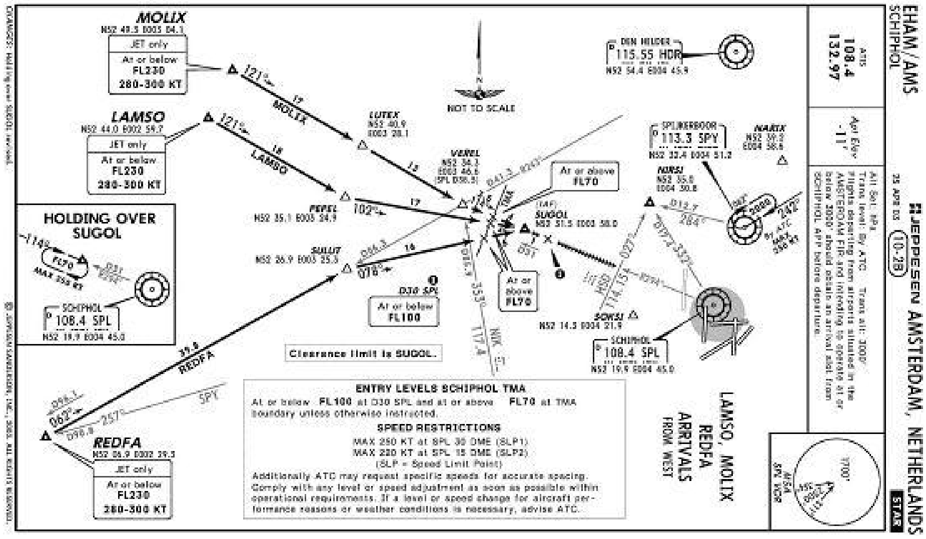


Figure 2: REDFA Standard Instrument Arrival Route (STAR) to Amsterdam Airport Schiphol[1]

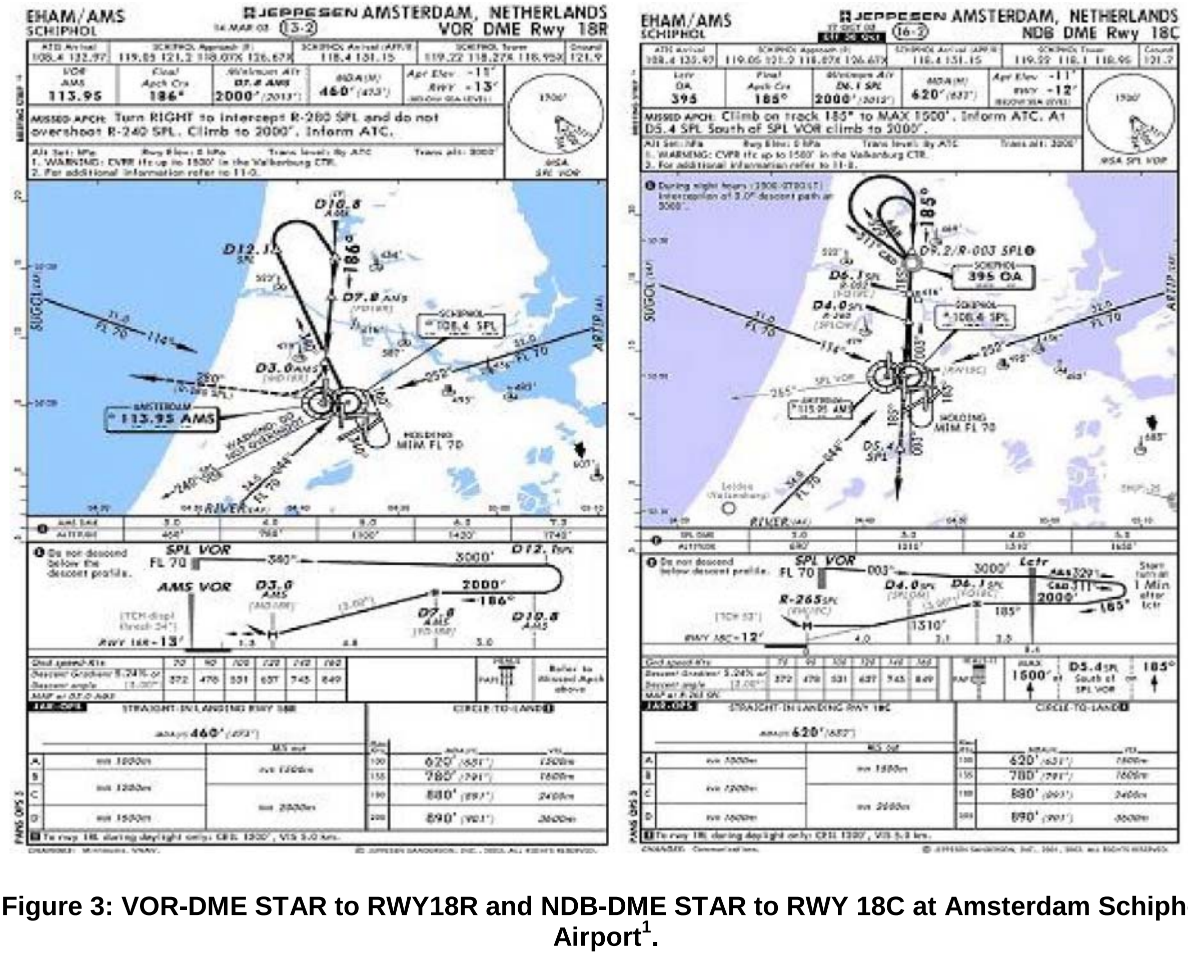


Figure 3: VOR-DME STAR to RWY18R and NDB-DME STAR to RWY 18C at Amsterdam Schiphol Airport[1].

[1] *Illustrations ©Jeppesen-Sanderson, Inc. Not to be used for navigation*

2.2.2. Flight phases and sub-phases

2.2.2.1. Preparation

Pilots were asked to complete the 'before operate' checklist just before the simulator was switched on at Flight Level (FL) 140. The sessions started with an ATC clearance to descend to FL70 and be level at 30nm from the SPL VOR. Pilots were asked to switch off the AutoPilot at SUGOL.

2.2.2.2. Descent

FL140 is referred to as T0. From there, it takes approximately 12 minutes to reach the final approach altitude of 2000ft.

Approaching FL100, the crews were expected to go through the approach checklist. When passing 30nm SPL, ATC issued radar vectors to position the aircraft on an intercept heading towards the final approach track. During this phase, the crew was instructed to disengage the autopilot and proceed by manual flight. The aircraft was configured for initial approach.

When passing FL100, ATC instructed the crew to decelerate to 250kias. The crews descended further to arrive at the final approach point at 2000ft and in the proper aircraft configuration. When the aircraft was established on the intercept course towards the runway, ATC cleared the crew for approach.

At approximately TO+12, the VOR needle became 'alive'. The aircraft was then configured for final approach (gear down and landing flap setting; speed approximately 150kias).

On final approach, from 2000ft to the ground, the PF had to perform a complex tracking task and the PNF had to manage communications, read the landing checklist and monitor (and coach) the PF and the progress of the approach. The final approach took about 2 minutes to fly.

2.2.2.3. Landing

Automatic GPWS call-outs were operational, and helped the aircrew to estimate when to start the flare. The cockpit height above the ground was in correspondence with the simulated aircraft type (a B747).

2.2.3. Crews

A total of four crews (paid volunteers) participated in the experiment. Standard operational procedures were kept simple to allow for slight differences between the activities of personnel from different companies. The flight experience of the pilots is presented in Table 1:

Table 1: Flight experience of the pilots

Pilot role	Commercial aircraft experience		B747 experience	
Crew 1, Captain	> 16 years	7300 hrs	--	-
Crew 1, F/O	1 year	650 hrs	1 year	650 hrs
Crew 2, Captain	7 years	3945 hrs	--	--
Crew 2, F/O	--	- (cpl)	--	--
Crew 3, Captain	1 ½ years	920 hrs	1 ½ years	920 hrs
Crew 3, F/O	--	- (cpl)	--	--
Crew 4, Captain	9 years	6500 hrs	½ year	4000 hrs
Crew 4, F/O	--	- (cpl)	--	--

The average age of the pilots was 29.5 years. As is shown in Table 1, Captains and First Officers (F/Os) differed considerably in their experience level. Also the overall experience with the B747 was limited. All pilots were familiar with Amsterdam Schiphol.

2.2.4. Crew Briefings

Prior to the trial, a briefing guide was sent to the crews. In the morning of the experiment, the crews received an oral briefing about the goals of the experiment, consisting of the following items:

- research briefing about the goals of this cognitive streaming trial
- general briefing about GRACE
- briefing on crew tasks and procedures

2.2.5. Training

Each crew received about 1 hour of training to familiarise them with the simulator. In addition this training served to harmonise the working procedures (crew members were not necessarily from the same operator).

2.2.6. Experimental Design

Based on the experience gained in the pilot study, it was decided that the flight crews would perform six runs per day. The first four runs would be balanced across teams in a Latin square arrangement, allowing full counter-balancing of the combinations of flying role and of control vs. party line conditions. The remaining two runs would be balanced as closely as possible, and would be used, if necessary, to repeat any of the first four runs during which data loss or other disruptions occurred.

Table 2. Experimental design. First character: C = Captain as PF; F = First Officer as PF. Second character: P = party line; N = no party line

Crew/Run	1	2	3	4	5	6
Crew 1	CN	CP	FN	FP	CN	FP
Crew 2	CP	FP	CN	FN	CP	FN
Crew 3	FN	CN	FP	CP	FP	CN
Crew 4	FP	FN	CP	CN	FP	CN

3. RESULTS

3.1. Methodology for the Analysis of Data per flight

This section describes the process of analysis, and how the data are presented. The flight is represented schematically in Figure 4.

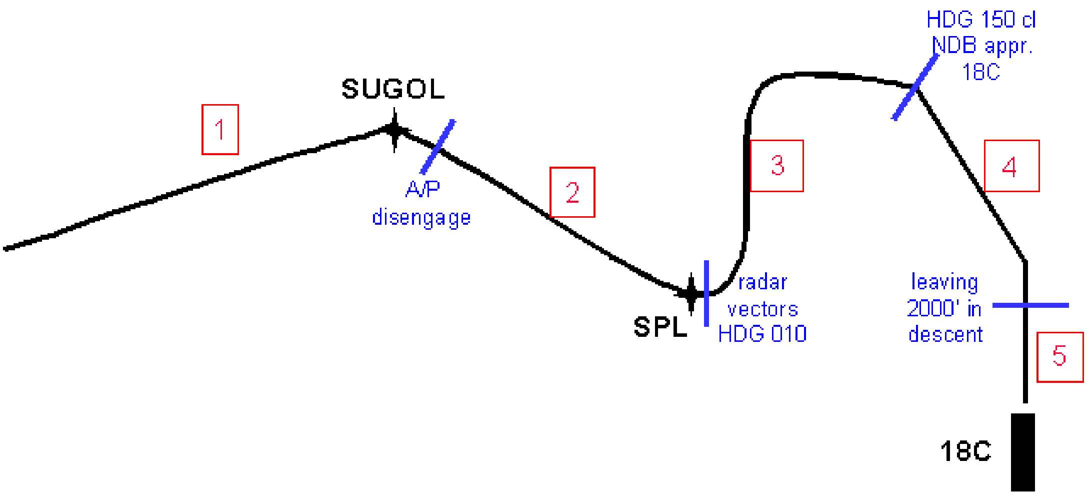


Figure 4: The five phases of flight as used during the data analysis

The first phase of the flight was a straight segment, flown on autopilot. After the SUGOL waypoint, the flight crew was asked to disengage the autopilot ("A/P disengage") and as a consequence the PF had to control the aircraft manually. A relatively long straight segment followed to the SPL waypoint and is depicted as Phase 2. Just after the SPL waypoint, the crew received the ATC instruction to change the heading to course 010. The small straight segment (Phase 3) was followed by a turn to the right to position the aircraft for the intercept of the extended runway centre line, The timing of the ATC 'turn-right' instruction varied due to other traffic. Phase 4 depicts the initial approach (flight segment intercepting the runway extended centre line), followed by Phase 5, the final approach and landing on runway 18C of Schiphol Airport.

For the analysis the following methodology was adopted:

- **Performance data**: the analysis of this kind of data requires relatively stable flight segments in order to be able to calculate the mean and standard deviation of the flight path deviations. Since the first segment was flown on autopilot, no human intervention occurs and performance data were not influenced by the R/T condition. Segment 3 was found to be too short for the performance analysis. Hence only segments 2, 4 and 5 are included in the analysis. As a specific measure, the accuracy of the touchdown location on the runway was also used.
- **Physiological**: physiological data analysis requires stable periods of similar subject activity levels. Inspection of the data led to the definition of five relatively short periods of 90 s to analyse the data for each flight phase (an exception being the EEG data, for which representative segments of 10 s were selected). Flight Phase 1 can in this respect be regarded as a base-line for the comparison of the activity levels in the subsequent phases. However, this does not apply to the R/T condition, since background R/T could also be present during Phase 1.
- **Subjective & observations**: Rating forms were completed after the completion of the whole flight. No separate forms were used during the flight, to reduce the confounding effect of interruptions by non-trial-related events. Therefore analysis results are available only per complete flight. The same applies to the data collected by the in-flight observer.
- **Debriefing**. All crews were debriefed at the end of the day after the whole experiment using a set of pre-defined questions.

The salient results of the analyses are described in the main part of this report. Details of the statistical analysis are given in Appendix E.

3.2. Performance data

3.2.1. Analysed data

Background R/T may influence performance data in the form of steadiness in maintaining predefined settings such as course or altitude. Since all crews used the flight director to help them to control the aircraft, steering commands of the flight director are selected as the main source of tracking information. Alternatively, the smoothness of the roll and pitch angles can also be examined as an indicator of stability of flight performance. Therefore both pitch and roll angle deviations, as identified by the standard deviation, were selected. Both the flight director commands and the steadiness of pitch and roll angles can be assessed only for known flight paths. In this experiment, only straight segments conform to this requirement. Therefore, for each of the straight segments of Flight Phases 2, 4 and 5, data periods were selected in which stable flight conditions occurred (= straight flight segment). The length of each analysis period was 60s. For each period, the following information was derived:

- Average and standard deviation of the flight director pitch and roll angle commands. Since the average command should be about zero (the command is related to the actual roll angle and required roll angle, and it is assumed that the pilot tries to minimise the commands), only the standard deviation is used in the statistical analysis. The average value was checked for the assumption of being close to zero, or, in other words, for the fact that the PF was following the steering information.
- The standard deviation of the pitch and roll angle, providing an estimate of the tracking error of the PF following a pre-set course and vertical speed. Note: none of the straight segments was flown at a fixed altitude due to the approach scenario.

The standard deviation of the heading angle (yaw) during the straight segments was considered for the analysis as an indication for the horizontal flight path tracking. However, the heading is the integral of the heading change over time, whilst the heading change is directly proportional to the roll angle via a smooth (non-linear) relation. Therefore analysing this parameter would not provide more information than the analysis of the roll angle, and as such the analysis of the heading angle was omitted.

Similar to the heading angle, vertical speed changes are directly proportional to aircraft pitch angle changes (provided that air speed and pressure remain approximately the same). The vertical speed indicator in aircraft is normally based on (small) air pressure changes, and due to the required mechanical construction lags considerably compared to the actual situation. Therefore pilots have been taught to use this indicator as a reference only, instead of as a flight control parameter. In addition, the vertical speed during the descents was not prescribed. Pilots selected values between 800 and 1000 ft/min based on the distance to the next waypoint after receiving ATC clearance to descend. Therefore it was decided not to analyse this parameter as an index of vertical path tracking.

Since during the final approach the flight path reference changed from following needles to using the outside vision to line up with the runway (both horizontally and vertically), the ground paths during the final approach are analysed in more detail.

Finally, the distance between the runway threshold and the aircraft touchdown point was looked at as a measure of how well the approach was performed. However, no instruction was given to the pilots that this measure was included. Also the length of runway 18C is more than sufficient to allow pilots to take a relaxed attitude to the actual touchdown point (they have more than ample space for the deceleration of the aircraft) in favour of landing smoothly.

In the figures below, the following abbreviations are used:

P: Party Line
NP: No Party Line
Cap: Captain
FO: First Officer

3.2.2. Performance results

3.2.3. Periods of Straight and Level Flight

3.2.3.1. Pitch angle deviations

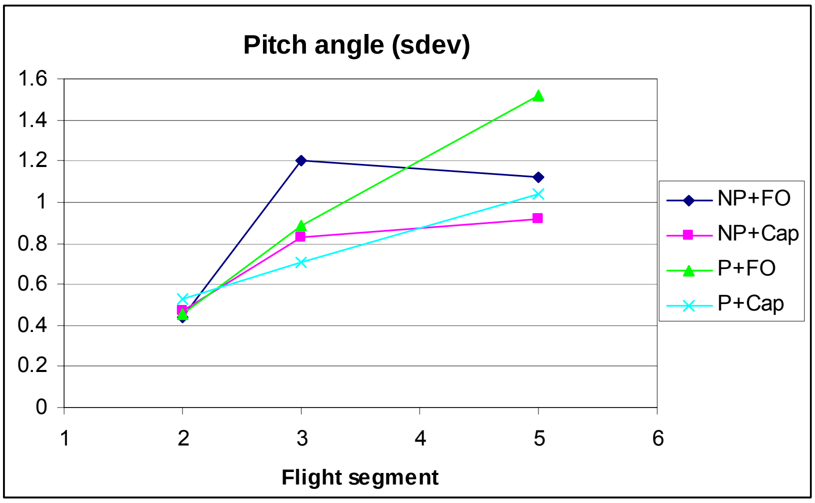


Figure 5: Standard deviation of the pitch angle versus stable periods in the flight phases

Figure 5 shows that the standard deviation of the pitch angle increases after Phase 2, meaning a less stabilised flight. No deterioration due to the party line R/T is present. The deviations for the FO-flying are larger than for the Captain, which can be explained by the difference in experience level.

3.2.3.2. Pitch command of the Flight Director

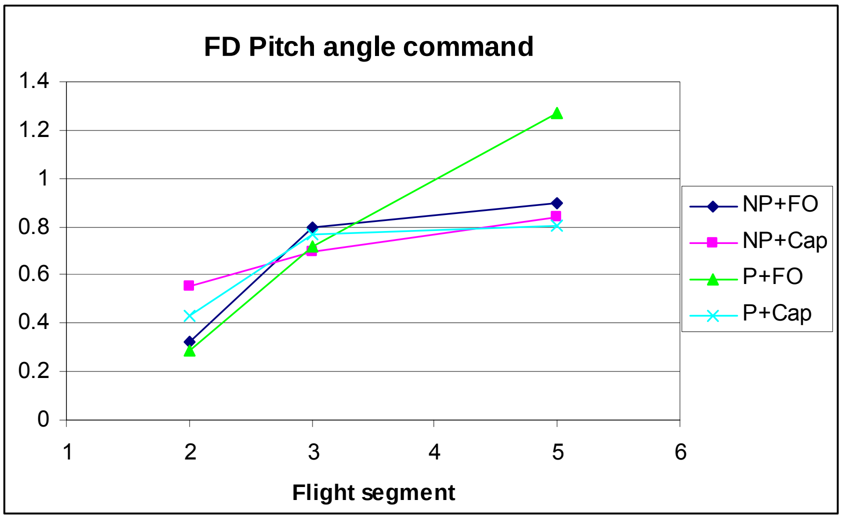


Figure 6: Standard deviation of the flight director pitch angle commands versus stable periods in the flight phases

The standard deviation of the flight director commands for the pitch angle provides a similar pattern as the standard deviation of the pitch angle. This is also an indication that the correct values were selected at the flight control panel and that the flight director information was used for primary flight control.

3.2.3.3. Roll angle deviations

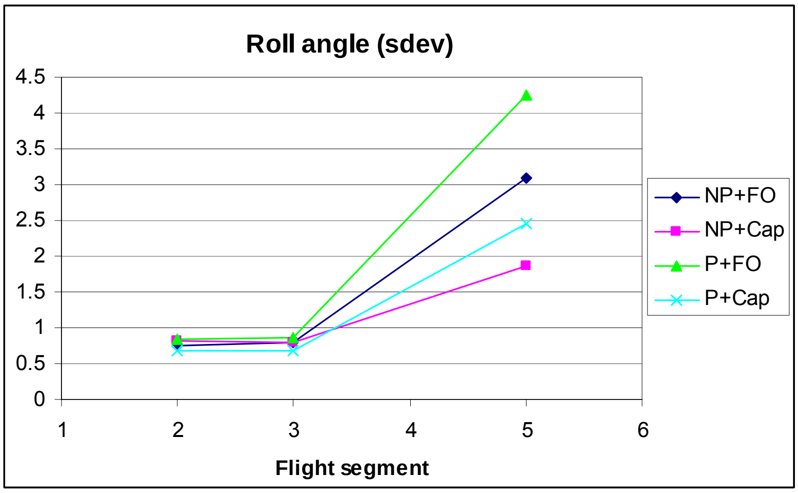


Figure 7: Standard deviation of the roll angle versus stable periods in the flight phases

Figure 6 shows a significant increase in the roll angle for Phase 5. This increase can be explained by the closeness of the runway: during the final approach the pilot had to line up the aircraft with the visual runway centre line. For this line-up, some heading changes and hence roll inputs were necessary. A difference between the captain and F/O is again apparent, which can be explained by the difference in experience level and a difference in the strategy adopted for the approach to the runway (Captains used the visual runway reference earlier than the F/Os). The party line effects are not clear (for F/Os, a decrease with party line on short final; for Captains, a slight increase).

3.2.3.4. Roll Bar deviation of the Flight Director

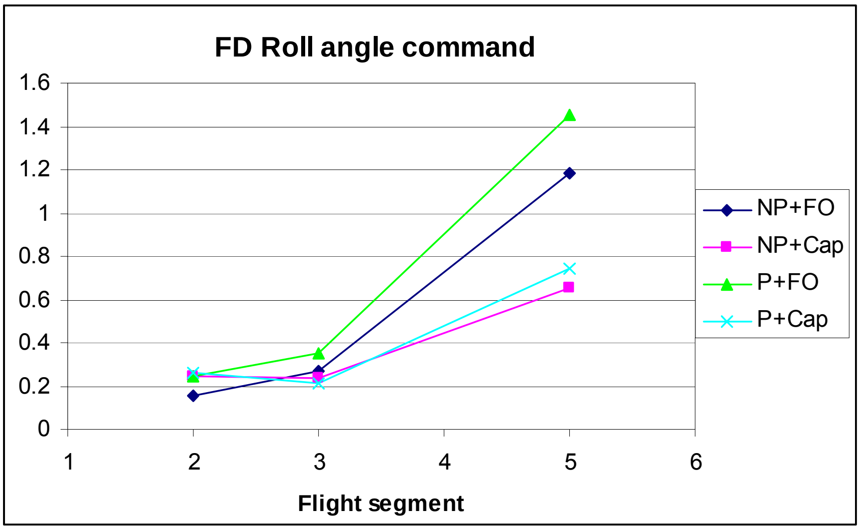


Figure 8: Standard deviation of the flight director roll angle commands versus stable periods in the flight phases

Analogous to the pitch angle, the standard deviation of the flight director commands for the roll angle follow closely the patterns as indicated in Figure 7.

3.2.3.5. Runway Proximity

A set of x, y, and z co-ordinates was recorded that showed the proximity of the aircraft to the runway. Example ground paths are shown in Figures 10 and 11.

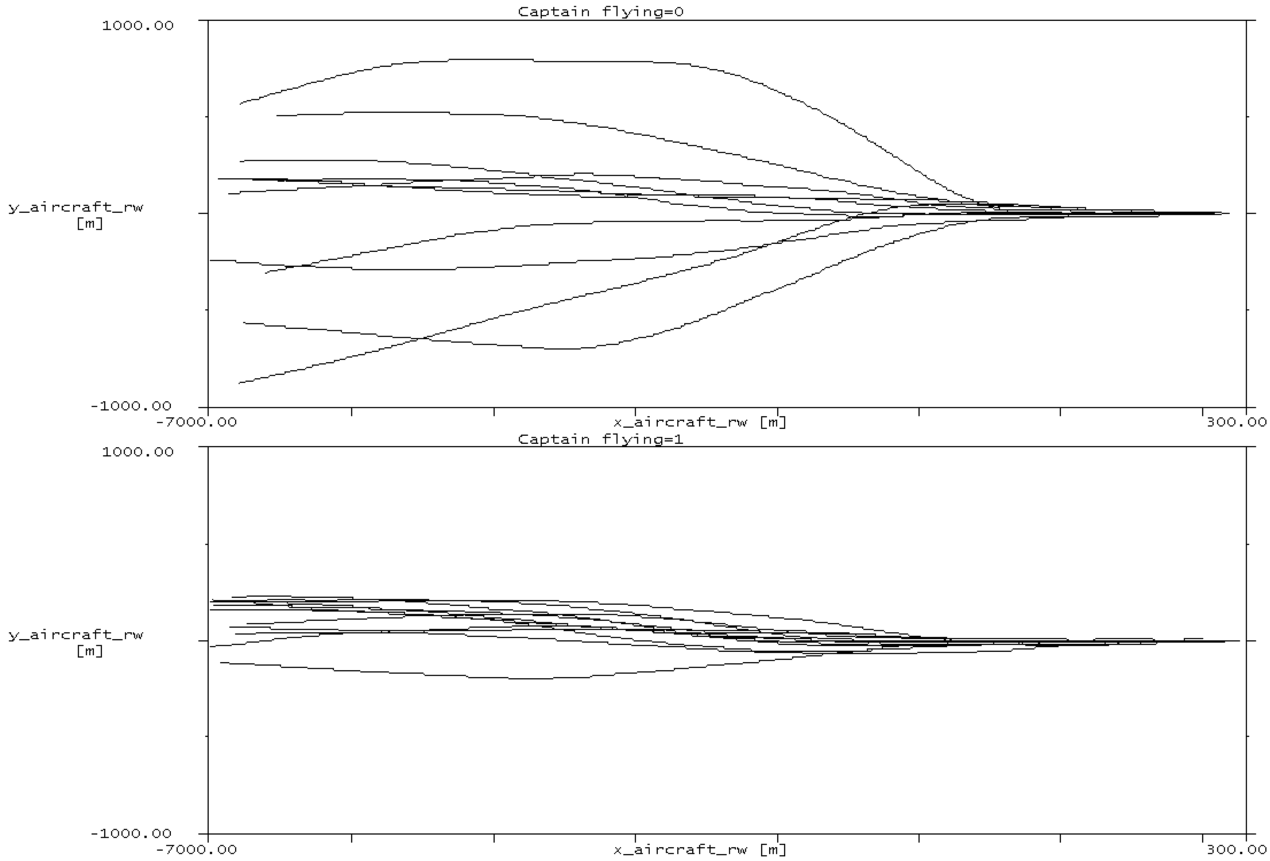


Figure 9: Horizontal ground paths of the approaches

Figure 9 presents the horizontal ground paths of the approaches. The runway threshold is at the right hand side of the picture. The upper part of the figure presents the ground paths of flights performed by the First Officer; the lower part were flights controlled by the Captain. The extended runway tracking accuracy for the initial approach differs considerably between the F/Os and Captains, but is corrected to similar values at short final. The conclusion that can be drawn from this figure is that the Captains used the visual runway information earlier in the flight to align with the extended runway centre line. The F/Os seem to follow the navigation instruments to a later stage in the approach. Since an NDB approach is relatively inaccurate, larger deviations result on relatively short final. Unfortunately, due to the large difference, it is no longer possible to investigate the effect of the party line R/T on the flight path accuracy.

The observed behaviour is also an explanation of the major differences found for the standard deviation of the roll-angle for Phase 5 in Figures 7 and 8.

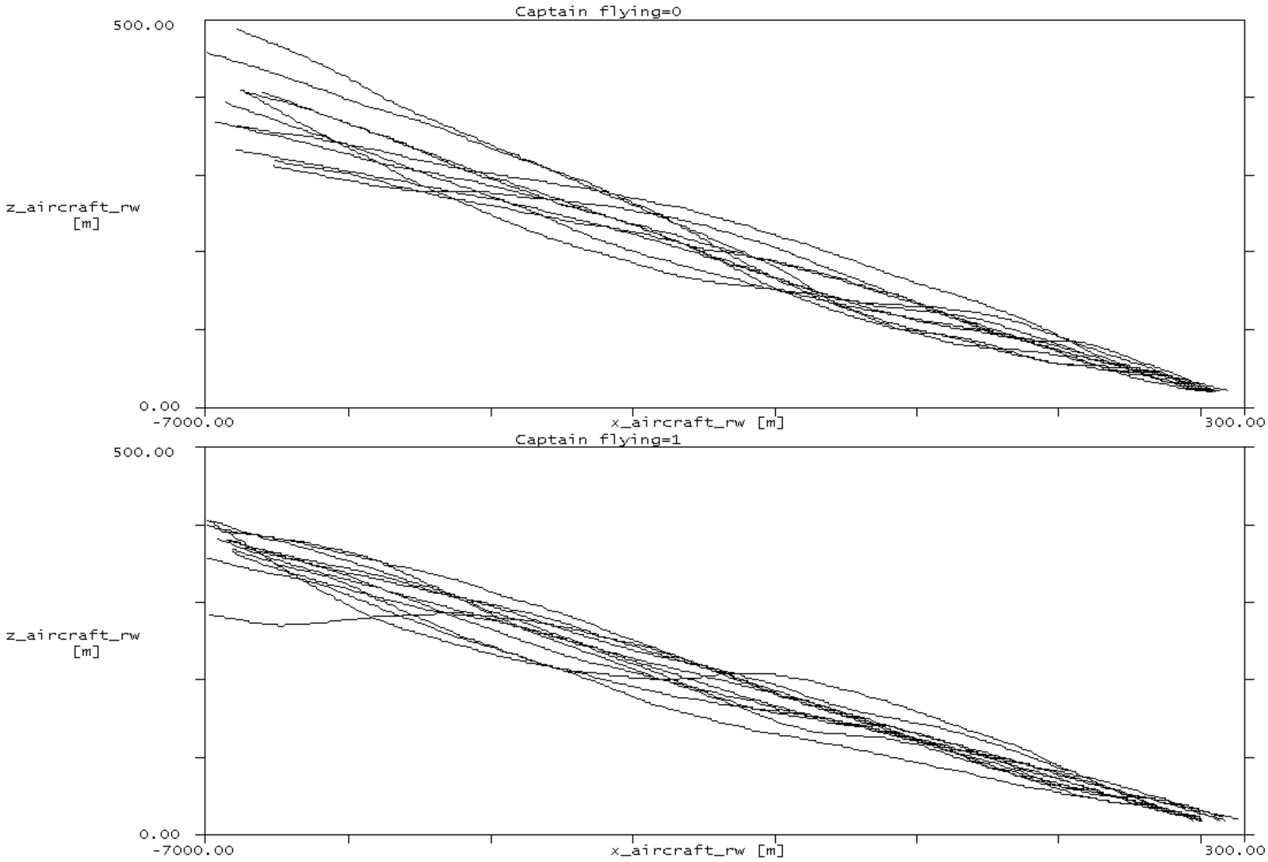


Figure 10: Vertical ground paths of the approaches

Figure 10 presents the vertical approach paths with the height above the runway (vertically) versus the distance to the runway threshold. Upper traces are from flights performed by the F/O; lower traces are from flights controlled by the Captain. The vertical flight paths are comparable for both the Captain and the F/O. Comparing the same paths for the influence of the party line R/T condition also revealed no visual differences.

3.2.3.6. Touchdown

The longitudinal and lateral co-ordinates of the touchdown point were recorded, allowing assessment of the accuracy of the landing ('touch-down point'). These co-ordinates were analysed using a repeated-measures multi-factorial multivariate ANOVA to determine the effect of the party line R/T and pilot role (Captain vs F/O). It was found that the party line condition was associated with a higher mean standard deviation of the longitudinal position on the runway if compared to the No Party Line condition.

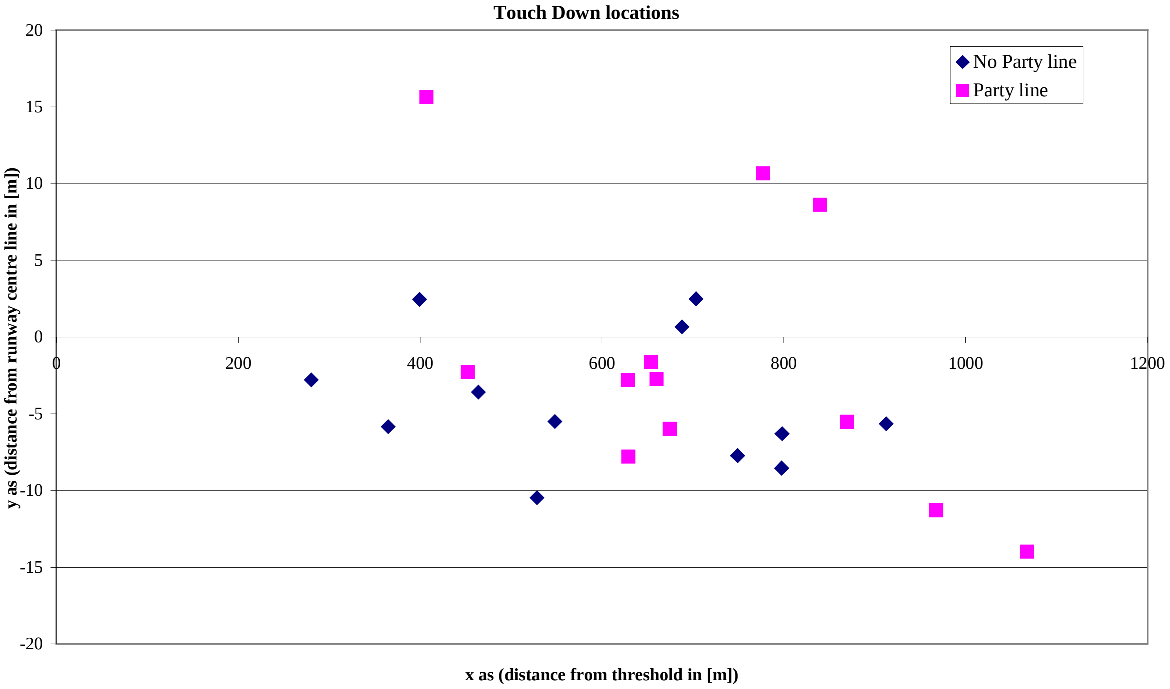


Figure 11: Footprint of the touch down positions on the runway. The runway threshold is at x-co-ordinate zero (left in the figure).

When looking at the landing performance between pilot role (Captain versus F/O) it was found that the landing accuracy of the F/O was lower than that of the Captain. Looking at the approach paths, this comes as no surprise. Statistically this effect was not significant.

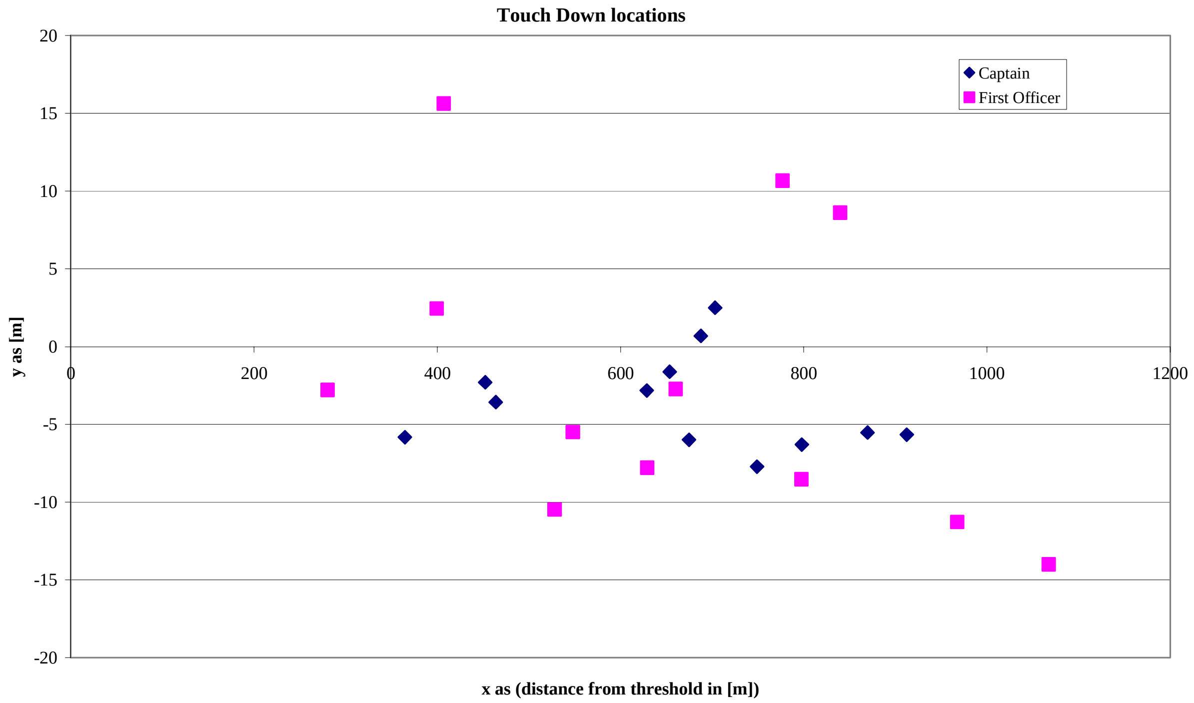


Figure 12: Touch down locations versus pilot role

3.3. Physiological data

Using NLR's Human factors Evaluations, data Analysis and Reduction Techniques (HEART) program, the physiological data were analysed in accordance with five phases of flight (Figure 4).

Data for captains and First Officers were analysed according to experimental condition; specifically, party line or no party line, and PF/PNF. Data from flights 5 and 6 were combined with data from the corresponding earlier experimental trial (CN, CP, FN or FP), to obtain an average value for that condition. Trial CS4FP1 was removed from analyses due to problems with that experimental run (wrong flight director mode selection).

All physiological data were synchronised off-line to the aircraft performance data using the Pseudo Random Noise bit code as transmitted by GRACE, with a resulting time-accuracy in the order of 15 ms (Hoogeboom, 2003).

3.3.1. Heart rate and heart rate variability

Heart rate was analysed by condition and flight phase (Figure 13). Per flight, five 90 s segments were identified as follows:
1. Starting 150 s before autopilot disengagement
2. Starting 10 s after autopilot disengagement
3. Starting whenever the altitude became less than 4800 ft
4. Starting 110 s before the altitude became less than 1750 ft
5. Starting 90 s before the altitude became less than 50 ft

Note: The selection was performed in such a way that each 90 s segment is representative of the flight phases as indicated in Figure 4.

A 4 (flight condition) x 5 (flight phase) repeated-measures ANOVA revealed a significant effect of flying condition. Pairwise comparisons showed that heart rate for the PF in the party line condition was significantly greater than heart rate in either of the PNF conditions. The same was true of the PF in the no party line condition: heart rate was much higher than that as PNF, either with or without the party line. A highly significant effect of flight phase was obtained. Further analysis revealed no significant difference between Phases 1 and 2, but significant differences between each of the other phases, such that heart rate increased throughout the duration of the flight. There was also a significant interaction between flight condition and flight phase. In Phases 1 and 2 there were no significant differences between flight conditions; however, in Phases 3, 4 and 5, heart rate in the PF party line condition was significantly greater than that of both PNF conditions. Additionally in Phase 5, heart rate in the PF no party line condition was significantly higher than that of the two PNF conditions.

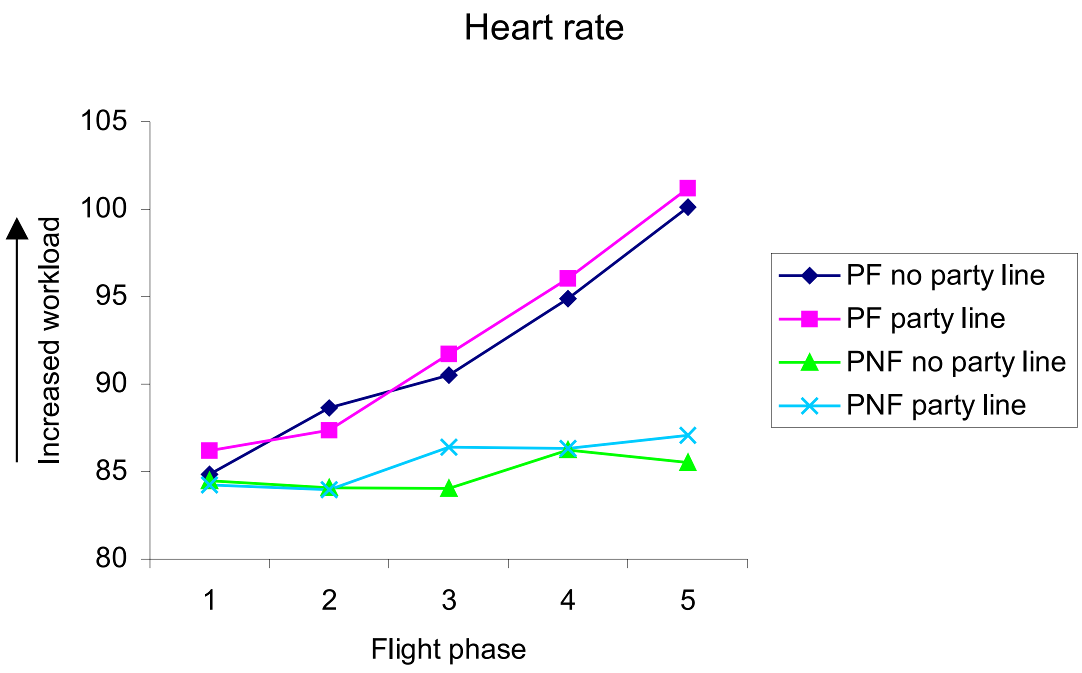


Figure 13: Average heart rate versus flight phases

In the PF party line condition there was no difference between Flight Phases 1 and 2, but heart rate increased steadily thereafter such that there were significant differences between each of the next two phases, and the difference between Phases 4 and 5 approached significance. For those acting as PF without party line, heart rate increased in a similar way throughout the flight. With the exception of Phases 2 and 3, there were significant differences between each of the flight phases, and the difference between Phases 1 and 2 approached significance. For the PNF conditions, the effect of flight phase was less apparent. With party line, heart rate in Phase 3 was significantly greater than it had been in either of the first two phases, but other comparisons – as well as those in the PNF no party line condition – did not reach significance.

Data for Captains and First Officers were analysed separately. There was a significant effect of flight condition for Captains, but not for First Officers. However, the main effect of flight phase observed in the overall data was true for both Captains and First Officers, as was the significant interaction.

Heart rate variability was also measured for each flight phase and flight condition (Figure 14).

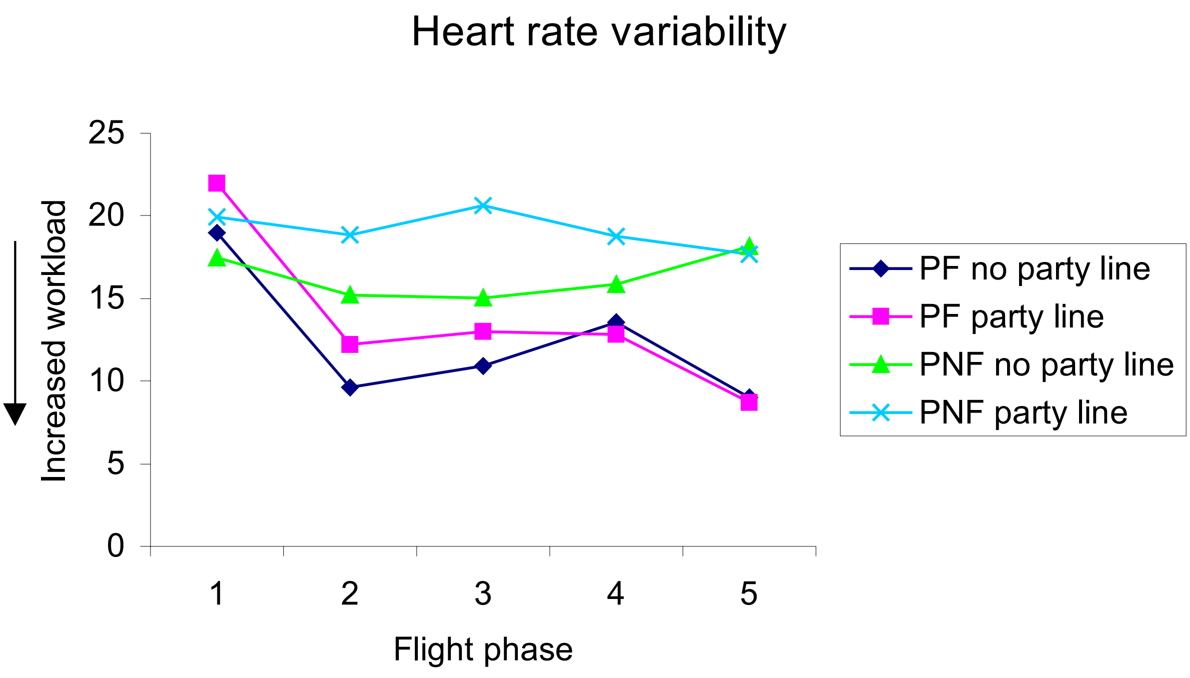


Figure 14: Heart rate variability (based on power in the 0.1 Hz frequency band) across flight phases, according to party line condition and flying role

A 4 (flight condition) x 5 (flight phase) repeated-measures ANOVA revealed a significant effect of flight condition, such that heart rate variability in the PF no party line condition was significantly lower than that of the two PNF conditions. There was a significant effect of flight phase: heart rate variability in Phase 1 was significantly greater than that in Phases 2, 4 and 5, and this difference approached significance in Phase 3. There was no significant interaction between condition and flight phase. When data were analysed separately in accordance with pilot rank, the effect of flight condition did not reach significance for either Captains or First Officers. There was a clear effect of flight phase for Captains, but this effect was not apparent in the data for First Officers. Consistent with the overall data, neither Captains nor First Officers showed an interactive effect of flight phase with condition.

3.3.2. EOG (blink rate)

Blink rate data were analysed by flight condition, and also by flight phases in accordance with the same 90s intervals used previously. The blink rate is derived from the vertical EOG channel, and is expected to decrease if the workload increases (Figure 15).

A 4 (flight condition) by 5 (flight phase) repeated-measures ANOVA revealed a significant effect of flight condition. Subsequent pairwise comparisons demonstrated that blink rate in the two PNF conditions was significantly higher than that observed in the two PF conditions. However, there was no significant difference between PF and PNF conditions in terms of the presence of the party line. There was also a significant main effect of flight phase: the first and last flight phases were significantly different from each of the intermediate phases, but not significantly different from each other. A significant interaction between flight condition and flight phase was observed. Pairwise comparisons showed:

PF versus PNF

- in Phase 3 there was no difference between the two PF conditions, but all other conditions were significantly different from each other
- in Phases 2 and 4, blink rate in the two PF conditions was significantly lower than in either of the two PNF conditions
- in Phase 5, the only difference was that blink rate in the PF party line condition was significantly lower than that of either of the two PNF conditions
- in Phase 1 there were no differences between PF and PNF conditions

Phases:

- in the PNF conditions, there were significant differences between Flight Phases 3 and 5, and 4 and 5, without party line; and Phases 2 and 5, and 4 and 5, with party line
- in the PF no party line condition, all phases were significantly different from each other except for Phases 1 and 2 in comparison with Phase 5
- there was also no significant difference between Phases 3 and 4
- in the PF party line condition, Phase 1 was significantly different from each of the other phases except the final one, and the final phase was different from Phase 4.

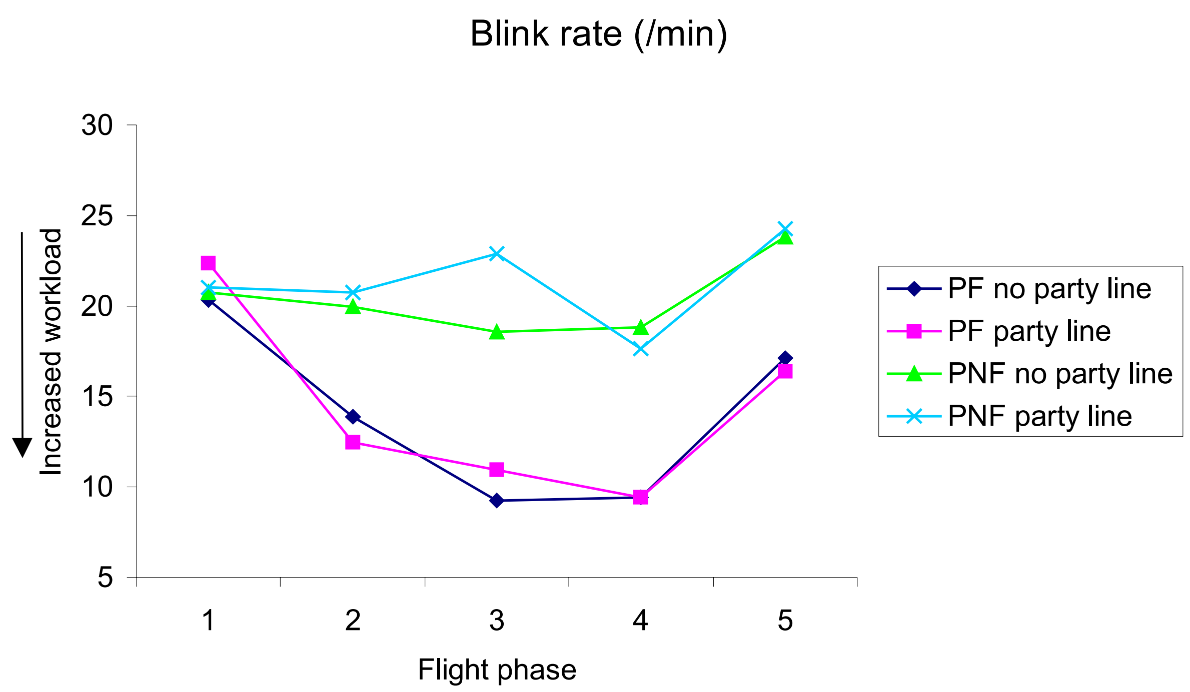


Figure 15: Mean eye blink rate (1/min) versus flight phases, according to party line condition and flying role

Data analysed separately according to pilot rank showed effects similar to the overall data. Both Captains and First Officers demonstrated significant effects of flying condition (blink rate in the two PF conditions being lower than in the two PNF conditions). First Officers also showed a significant effect of flight phase and an interaction between flight phase and condition. For Captains, the main effect of flight phase did not quite reach significance, nor did the flight phase/condition interaction.

Summary: Blink rate was significantly lower in the two PF conditions than in the two PNF conditions, especially during the intermediate stages of flight. This indicates that, as might be expected, the visual workload for the PF was higher than that of the PNF and that this workload increased after the autopilot was turned off.

3.3.3. EEG

Please see Appendix D for analysis and interpretation of the EEG data.

3.3.4. $SP0_2$ (Captains only)

$SP0_2$ shows the satiation of haemoglobin in the blood with oxygen, which *decreases* with high physical effort or in extreme situations. One-minute intervals were identified at each flight phase, and a mean percentage level was calculated for each phase according to flight condition (Figure 16).

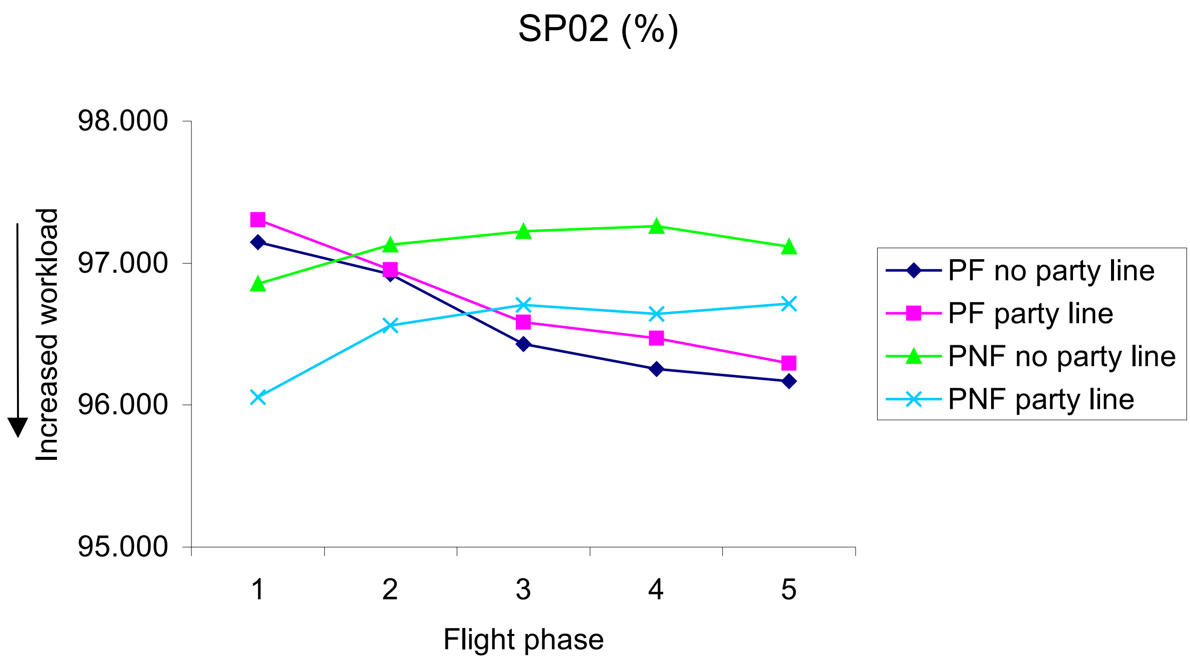


Figure 16: $SP0_2$ across phases of flight, according to party line condition and flying role

A 5 (flight phase) x 4 (flight condition) repeated-measures ANOVA was conducted on the data obtained from the four captains, revealing no significant main effects of flight condition or of flight phase, although there was a significant interaction. Further analyses (least significant difference [LSD] simple effects) showed that, for Phase 1, $SP0_2$ levels were significantly lower in the PNF party line condition than in either of the PF conditions. In the PF party line condition, there was a significant difference between Phases 1 and 5; and in the PNF no party line condition there was a significant difference between Phases 2 and 5.

3.3.5. Electrodermal activity (First officers only)

EDA levels were obtained for each flight phase and each flight condition (Figure 17).

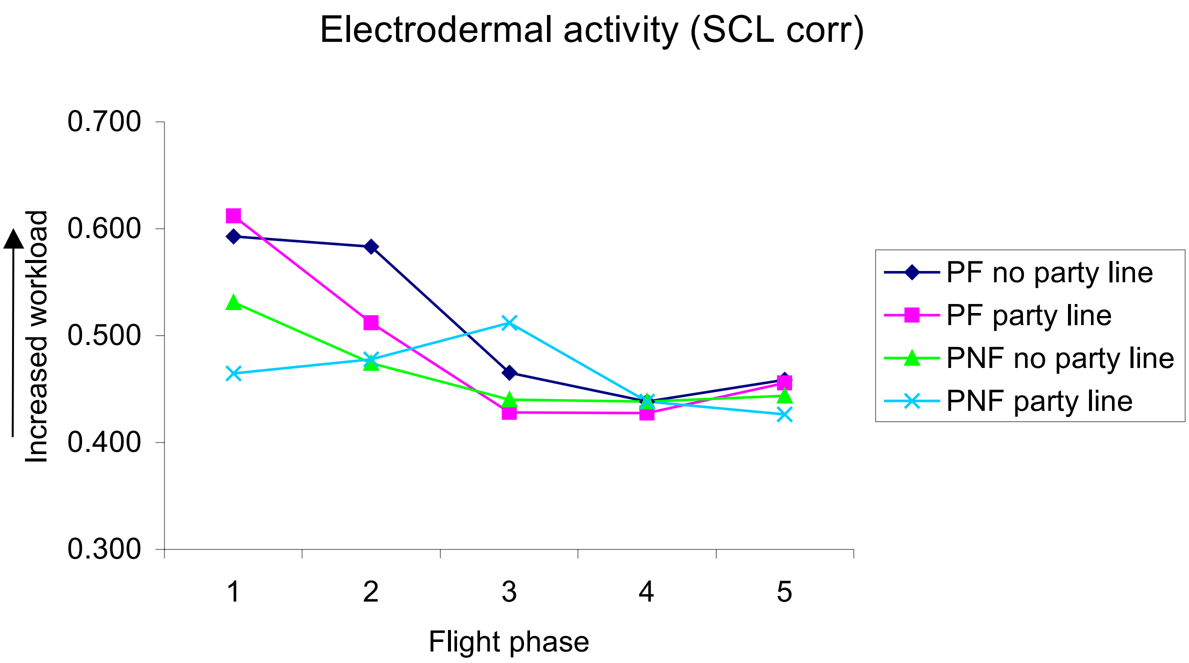


Figure 17: Electrodermal activity (skin conductance level) across phases of flight, according to party line condition and flying role

One-minute intervals were taken and SCL levels (mean values) were corrected according to a procedure by Lykken et al (1966). A 5 (flight phase) x 4 (flight condition) repeated-measures ANOVA showed no main effect of flight condition. There was, however, a main effect of flight phase. Pairwise comparisons showed that skin conductance levels in Phases 1 and 2 were higher than those in Phases 3, 4 and 5. There was no significant interaction between condition and flight phase.

The standard deviation of the SCL value during each flight phase was also analysed (Figure 18). This variation is a rough estimate of spontaneous fluctuations, and can be seen as another indicator of high activation. A 4 (flight condition) x 5 (flight phase) repeated-measures ANOVA showed no main effect of flight condition, although there was a significant effect of flight phase: the first phase was significantly different from Phases 2, 3 and 4. There was also a significant interaction between condition and phase. Pairwise comparisons showed that SCL deviations in Phase 3 were significantly higher in the PNF party line condition than the PF no party line condition; this pattern was reversed in Phase 5, where the PNF party line condition was significantly lower than the PF no party line condition. The difference between PF and PNF party line conditions approached significance at Phase 4, with the SCL deviation of the PNF being greater than that of the PF. In the PF no party line condition, SCL deviations were significantly higher in the first and last phases compared to deviations in each of the intermediate phases. In the PF party line condition, SCL deviation was significantly higher in Phase 1 than in Phases 3 and 4.

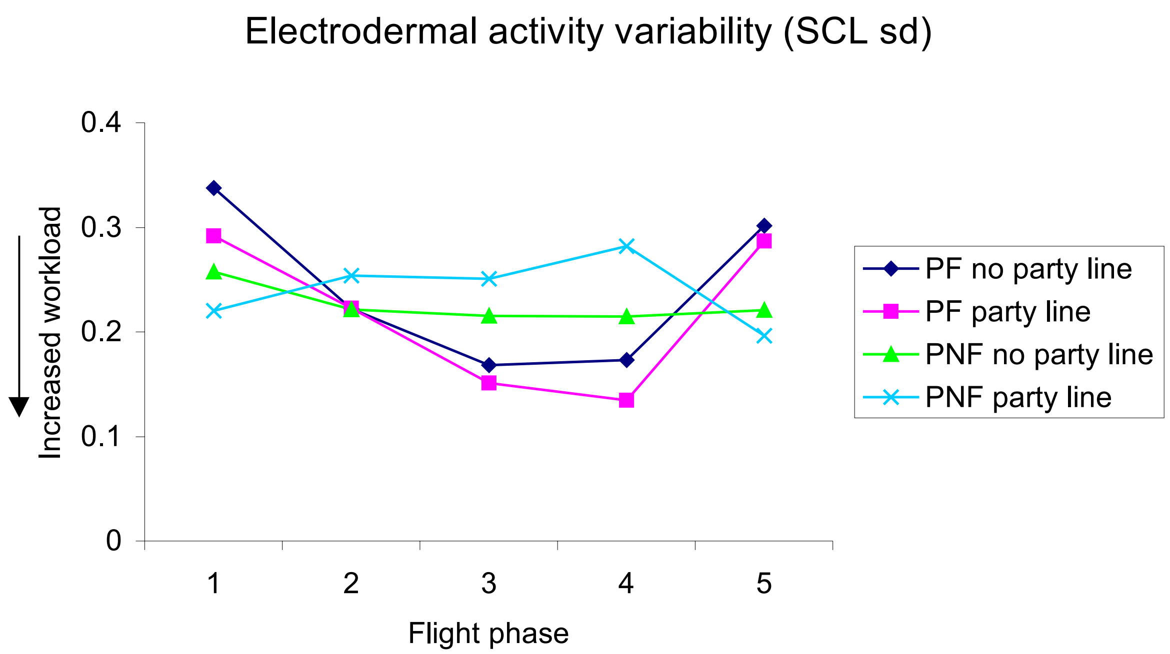


Figure 18: Standard deviation of SCL across phases of flight, according to party line condition and flying role

3.4. Questionnaires

Post flight questionnaires (in English) were administered to both crew members immediately after each landing. The Rating Scale Mental Effort (RSME) required pilots to place a line or cross at the appropriate point on a vertical scale of 0 to 150 to indicate the level of mental effort invested in the execution of the previous flight task. To help illustrate the use of the scale, nine verbal labels were also provided adjacent to the axis, ranging from ‘Absolutely no effort’ to ‘Extreme effort’ (see also Appendix B).

Pilots also completed an adapted version of the NASA TLX, which comprised eight questions referring to specific aspects of the flight, such as mental and perceptual activity, time pressure, and workload, as well as the level of distraction caused by R/T communications and perceived flight safety. A horizontal scale accompanied each item, ranging from 0 to 100, the intersection of which indicated the participants’ subjective response to each question. See Appendix B for the form used.

3.4.1. Statistical analysis procedure

Each crew performed a total of six landings, three with the Captain as PF and three with the First Officer as PF. In three of the flights, the crew received background R/T in addition to the normal R/T for their aircraft. These instances (Captain or First Officer as PF, and party line or no party line R/T) were counterbalanced across crews for the first four scenarios in a partial Latin square design (Table 2). In the remaining two scenarios (5 and 6), the constraints were that a) the Captain performed half of the flights, and b) half were performed in the presence of the party line. To analyse the questionnaires across each of the four conditions, the ratings given after Runs 5 and 6 were combined with those given following the corresponding earlier trials, and an average rating obtained. Due to problems with the first experimental run performed by Crew 4 in which the F/O made a wrong system selection, this trial was removed from analyses and only data from the corresponding condition in Run 5 were used.

Data for Captains and First Officers were combined, and a 2 (party line condition) x 2 (pilot flying condition) repeated-measures ANOVA was performed on the scores obtained for each questionnaire item. Responses were also analysed separately for both Captains and First Officers. Due to the limited number of trials, it was not possible formally to verify the normal distribution of the data (only four data points per condition in the 2 x 2 design due to the combination of the ‘double runs’). As usual, it was assumed that the rating scales were interpreted linearly by the participants.

3.4.2. RSME

Pilots indicated after each scenario the level of mental effort that had been required to perform the preceding flight task (Figure 12). A 2 x 2 repeated-measures ANOVA revealed a significant main effect of party line condition, such that crews rated those flights with background R/T as requiring a greater level of effort to perform than those in the silent cockpit condition. There was also a significant effect of flying role: pilots reported the need for a greater level of mental effort when acting as PF than as PNF. No significant interaction was found between party line condition and pilot flying role.

Data were also analysed separately for Captains and First Officers. The effect of party line condition did not reach significance for First Officers, but Captains reported that the flights with background R/T required a greater level of mental effort to perform. Although the overall data showed a significant difference in the rating of mental effort required for the PF compared to the PNF, this did not reach significance for either Captains or First Officers alone. Consistent with the overall data, neither Captains nor First Officers showed an interaction between party line condition and flying role in their ratings of mental effort.

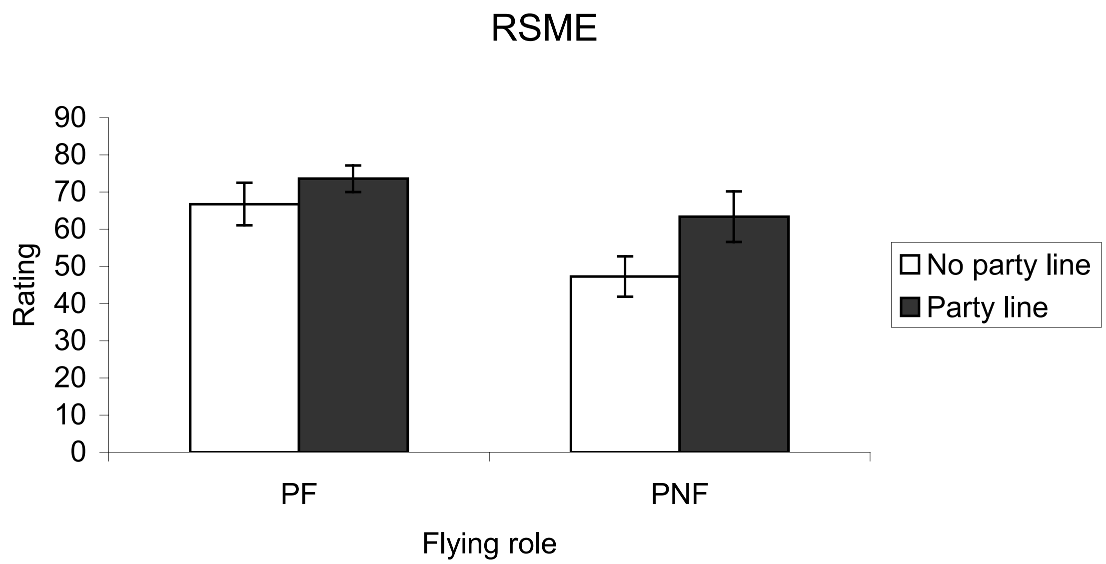


Figure 19: RSME score in party line and flying conditions, for PF and PNF. Error bars show standard error. Note: a rating of 70 corresponds to 'considerable effort'

3.4.3. Mental and perceptual activity

Pilots were asked to rate how much mental and perceptual activity (thinking, deciding, calculating, remembering) was required to perform each flight task (Figure 20). A 2 x 2 repeated measures ANOVA revealed a significant effect of party line condition, indicating that more mental and perceptual activity was required in the presence of background R/T than without. There was also a significant effect of flying role, such that more mental and perceptual activity was required as PF than as PNF. Furthermore, a near-significant interaction was obtained between flying role and party line condition, indicating that the increase in mental and perceptual activity imposed by the presence of background R/T was greater for the PNF than the PF.

The effect of party line condition on ratings of mental and perceptual activity reached significance for the Captains alone. Although significant in the overall data, the reduced power meant that neither Captains nor First Officers exhibited a main effect of pilot flying role. The Captains' ratings showed no interaction between party line and flying role, although for First Officers this approached significance. As with the overall data, this interaction indicated that the increase in mental activity due to the party line was greater for the PNF than the PF.

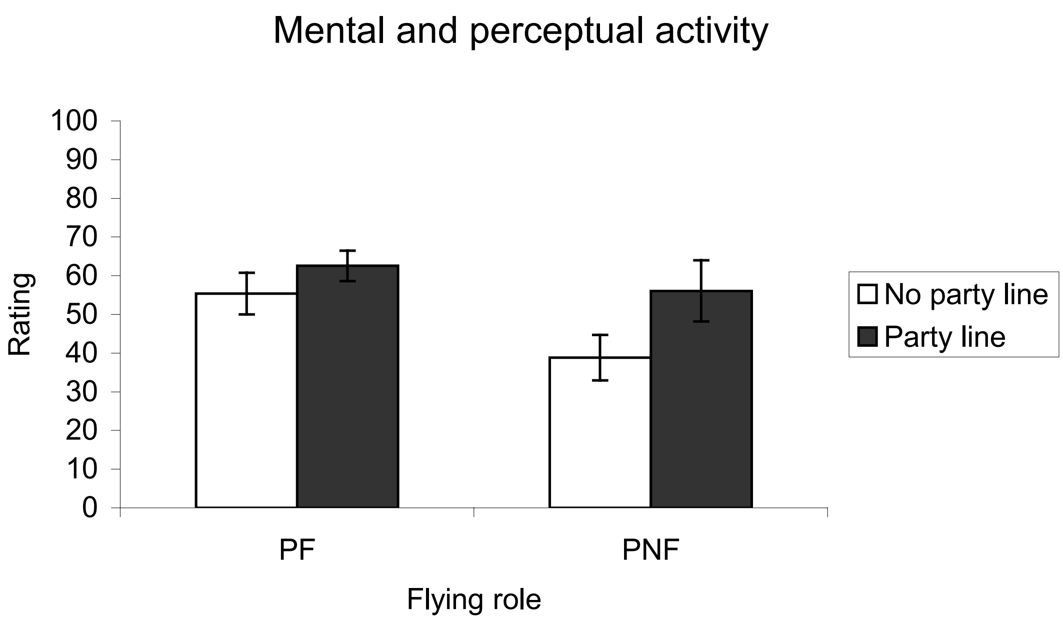


Figure 20: Percentage rating of mental and perceptual activity according to party line condition and flying condition. Error bars show standard error. A rating of 0 corresponds to 'None'; a rating of 100 indicates a high activity level (but no absolute value as compared to the RSME scale)

3.4.4. Time pressure

Crew members were asked to indicate how much time pressure they felt due to the rate or pace at which the task elements occurred (Figure 21). A 2 x 2 repeated measures ANOVA showed a significant main effect of party line condition, such that perceived time pressure was increased in the presence of background R/T. Also, a significant effect of flying role was obtained, with time pressure perceived as greater for the PF than the PNF. The interaction between party line and pilot flying role approached significance, with the increase in time pressure due to the party line being greater for the PNF than the PF.

When data from the Captains and First Officers were analysed separately, Captains showed a significant main effect of party line on perceived time pressure but First Officers did not. On the other hand, First Officers demonstrated a significant main effect of flying role — such that time pressure was greater for the PF than the PNF — but Captains did not. Ratings from First Officers also showed a significant interaction, indicating that the increase in time pressure caused by the party line was greater for the PNF than the PF. Captains did not show this pattern.

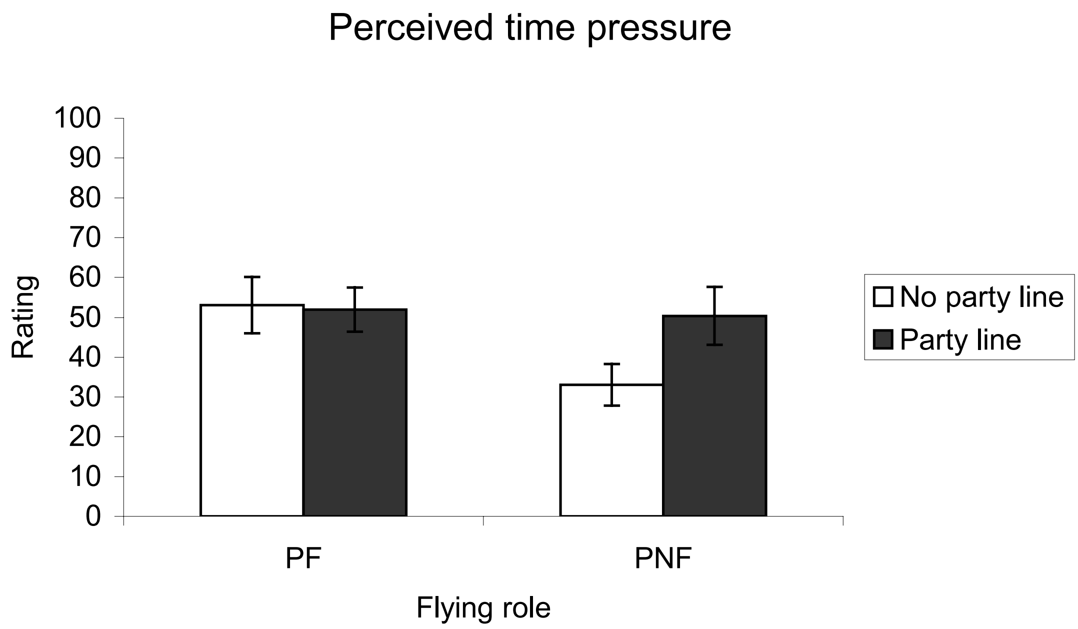


Figure 21: Percentage rating of perceived time pressure according to party line condition and flying condition. Error bars show standard error. A rating of 0 is assumed to be 'slow and leisurely' whereas a rating of 100 would represent 'rapid and frantic'

3.4.5. Success in accomplishing tasks

Pilots were asked to indicate how successful they believed they were in accomplishing the tasks required during the flight task (Figure 22). A 2 x 2 repeated measures ANOVA revealed no significant main effects of party line or flying role. Also, there was no significant interaction between party line and flying role.

When responses from Captains and First Officers were considered separately, Captains showed a near-significant effect of party line, indicating that they felt more successful at accomplishing their tasks in the silent cockpit condition. This was not the case for First Officers, who showed no effect of party line on perceived task success. Neither Captains nor First Officers demonstrated an effect of flying role or any interaction between flying role and party line condition.

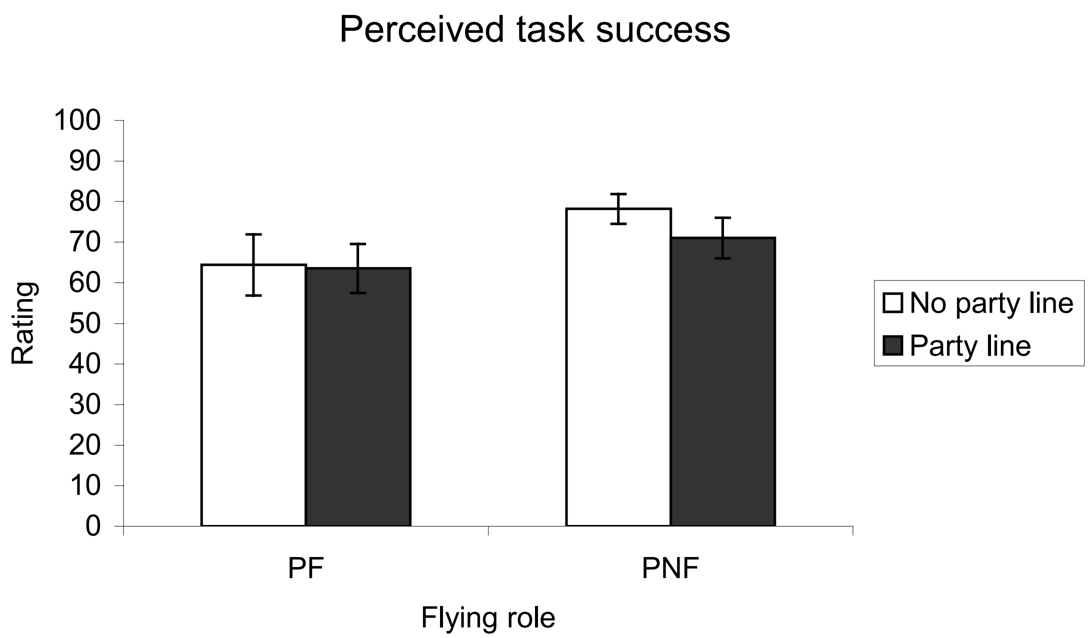


Figure 22: Percentage rating of 'perceived success in accomplishing tasks' according to party line condition and flying condition. Error bars show standard error. A rating of 0 indicates 'no success' whereas a rating of 100 indicates 'full success'

3.4.6. Mental workload

Pilots indicated on a scale of 0 to 100 how hard they felt they had to work (mentally) to accomplish their level of performance (Figure 16). There was a significant effect of party line, such that pilots felt they had to work harder in the presence of background R/T. A significant effect of flying role was also obtained, pilots indicating that the mental workload was greater for the PF than the PNF role. There was no significant interaction between party line and flying role.

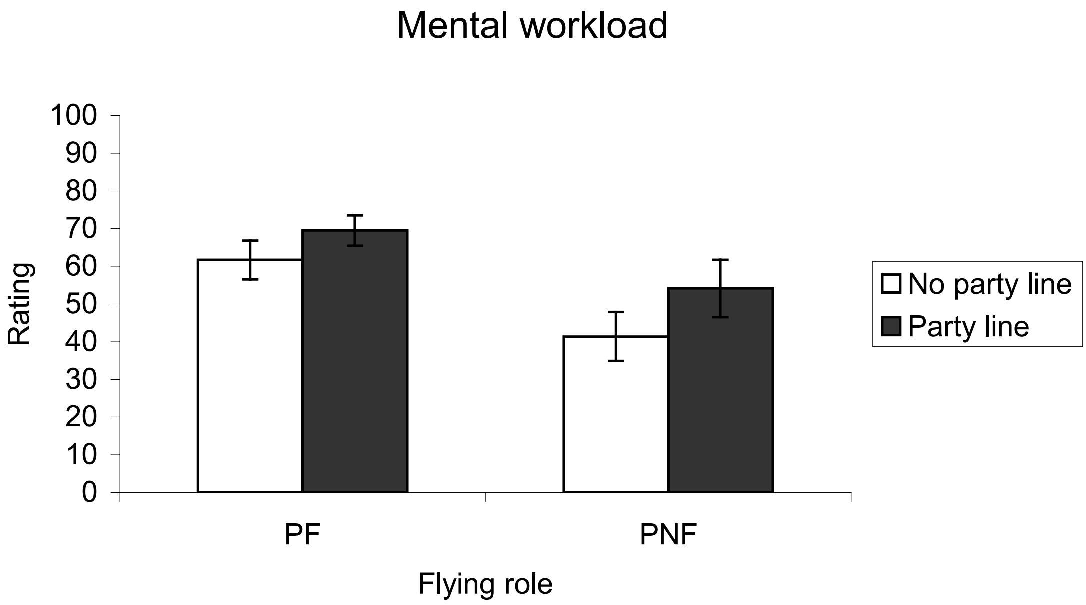


Figure 23: Percentage rating of 'mental workload' according to party line condition and flying condition. Error bars show standard error. A rating of 0 would indicate no effort, whereas a rating of 100 would indicate a high level of workload. Note: the above figure can be compared with the RSME results, which use a more absolute scale

The effect of party line was significant for both Captains and First Officers, indicating that background R/T increased mental workload. Both Captains and First Officers demonstrated a significant effect of flying role, rating mental workload as higher for the PF than the PNF. For First Officers, there was also a significant interaction, such that ratings of mental workload were greater for the PNF than the PF role in the presence of background R/T. This effect was absent from the data for Captains.

3.4.7. Frustration

Pilots were asked to rate how insecure, discouraged, irritated, stressed and annoyed (versus secure, gratified, content, relaxed and complacent) they felt during the flight task (Figure 24). A 2 x 2 repeated-measures ANOVA revealed an effect of party line, frustration increasing significantly in the presence of background R/T. Pilots' ratings of frustration were not influenced by flying role, and there was also no significant interaction.

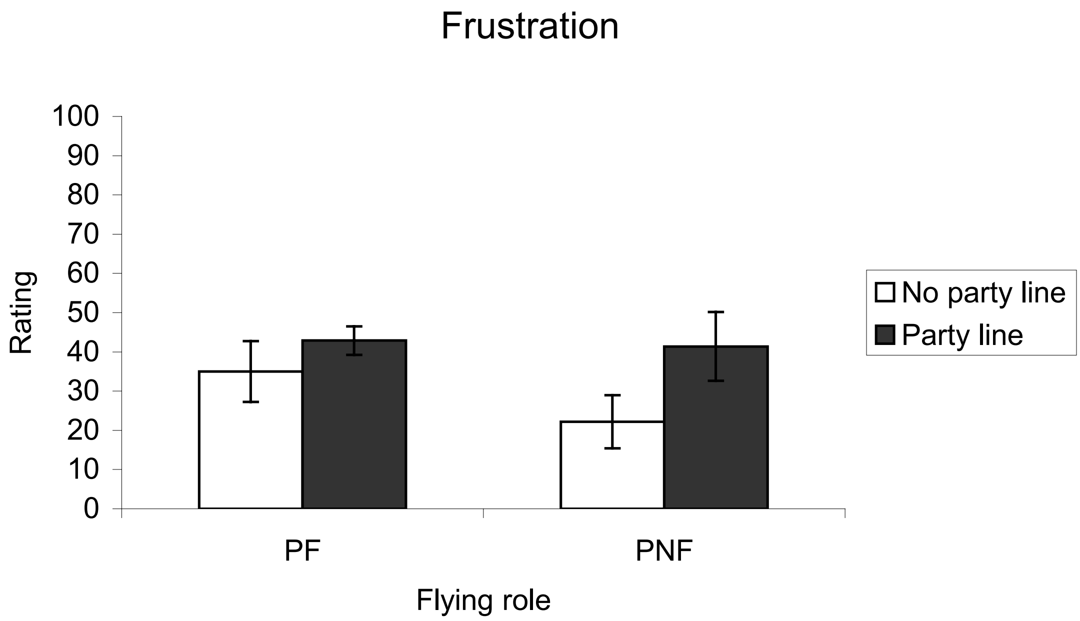


Figure 24: Percentage rating of 'frustration' according to party line condition and flying condition. Error bars show standard error. A rating of 0 denotes 'secure, gratified, content, relaxed and complacent', whereas a rating of 100 corresponds to 'insecure, discouraged, irritated, stressed and annoyed'

There was a significant effect of party line on frustration for Captains, but this effect did not reach significance for First Officers. Consistent with the results of the overall data, there was no effect of flying role on stress and frustration for either Captains or First Officers. Similarly, neither Captains nor First Officers showed an interaction between flying role and party line in their ratings of frustration.

3.4.8. Distraction

Pilots were required to rate the extent to which they felt distracted by R/T communications during their performance of the flight task (Figure 18). As expected, there was a significant effect of party line, pilots being more distracted in the condition with background R/T compared to that of the condition with own R/T only. The effect of flying role approached significance, pilots suffering more distraction from the R/T as PNF than as PF. There was no significant interaction between party line and flying role.

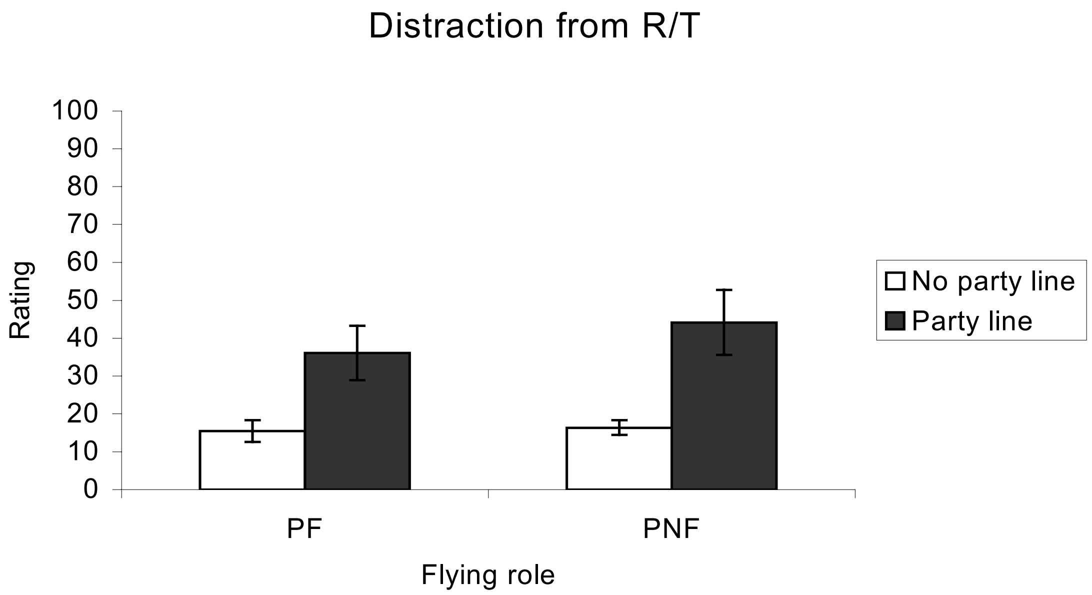


Figure 25: Per*centage rating of 'distraction caused by R/T communication' according to party line condition and flying condition. Error bars show standard error. A score of 0 indicates no distraction whereas 100 indicates complete distraction*

Although Captains showed no main effect of party line condition, First Officers were more likely to report that they found the background R/T distracting. Unlike Captains, First Officers were also more likely to report that they were more distracted by the R/T as PNF than PF, as shown by a near-significant effect of flying role. No significant interactions were obtained for either pilot rank.

3.4.9. Performance impairment

Pilots rated the extent to which they thought R/T communication had impaired their performance on the flight task (Figure 26). There was a significant main effect of party line; pilots believed that their performance was impaired by R/T communications to a greater extent in the presence of party line. There was no effect of flying role, i.e., pilots believed that R/T affected their performance to the same extent in the PF and PNF roles. There was no significant interaction between party line and flying role.

Although the effect of party line did not reach significance for Captains, First Officers were more likely to rate their performance as being impaired by R/T in the party line condition. There was no difference in ratings of impairment when PF or PNF for either Captains or First Officers. Also, in line with the overall data, neither Captains nor First Officers demonstrated an interaction between flying role and party line.

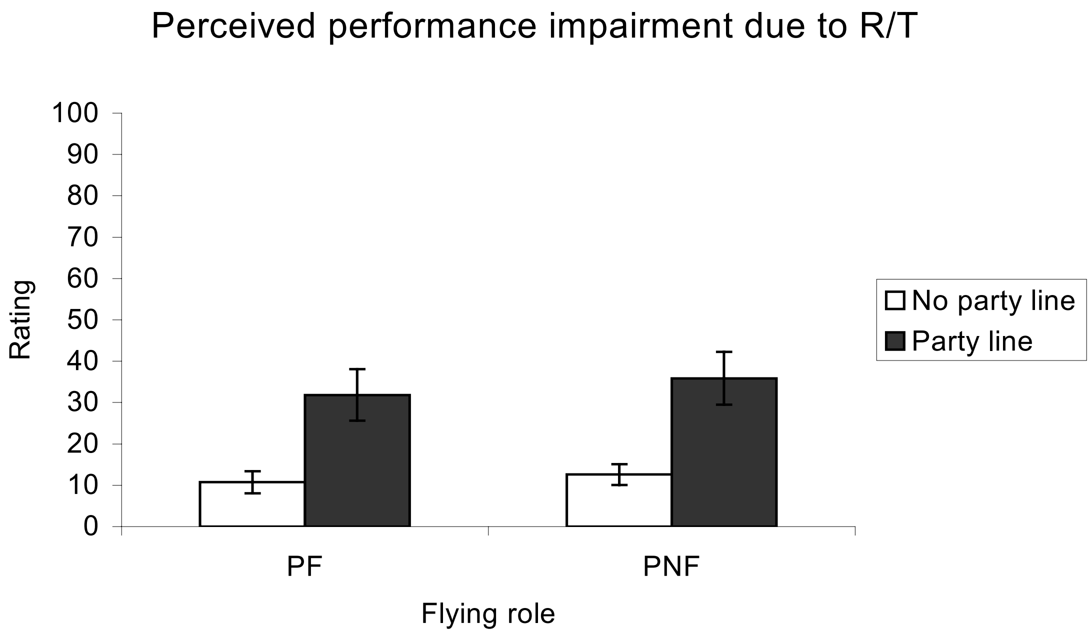


Figure 26: Percentage rating of 'perceived performance impairment' caused by R/T communication, according to party line condition and flying condition. Error bars show standard error. A value of 0 indicates no perceived performance impairment, whereas a value of 100 denotes strong impairment

3.4.10. Perceived flight safety

Pilots were required to rate the extent to which they felt that the flight that they had performed was safe (Figure 27). A significant effect of party line was obtained, such that perceived flight safety was decreased with background R/T. There was no significant effect of flying role, indicating that perceived flight safety was similar for PF and PNF roles. There was no interaction between party line and flying role.

Results analysed separately for Captains and First Officers showed similar effects. Both showed significant effects of party line, such that they perceived flight safety as being slightly lower in the conditions with background R/T. Perceived flight safety was not affected by flying role for either Captains or First Officers, Also, neither showed a significant interaction between party line and flying role.

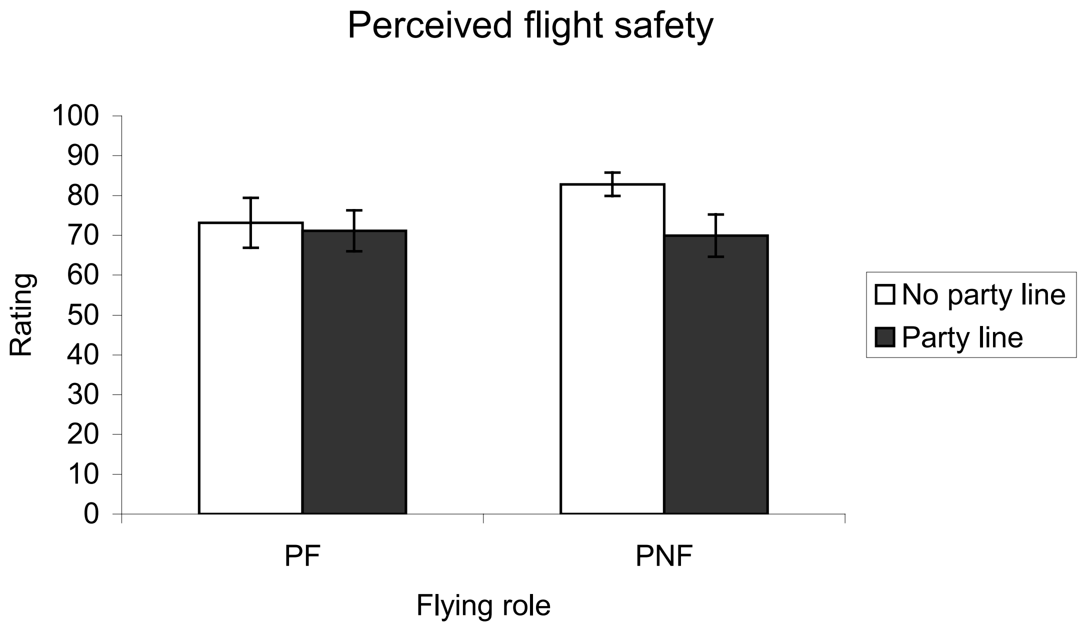


Figure 27: Percentage rating of 'perceived flight safety' according to party line condition and flying condition. Error bars show standard error. A value of 0 would indicate an unsafe flight whereas 100 would indicate a safe flight

3.5. Observational data

In-flight observations were made, noting instances of pilot errors concerning ATC calls and the completion of checklist items (see Appendix F for details). As usual, it was found that the absolute number of errors made was rather low; therefore, no statistical analysis could be performed. However, it is striking to see that the number of communication errors (missed calls and incorrect read-backs) almost doubled in the party line situation, whilst the number of queried calls (pilot request confirmation) remained almost equal.

When looking at the instances of checklist item errors, no trend with respect to the presence of the party line is visible. Also the number of mistakes is relatively low. However, looking at the influence of ATC clearances versus the missed checklist items, the number of mistakes made seems to be higher. A trend emerges that RT *interruptions* may affect the effectiveness of flight deck procedures, a phenomenon in line with existing cognitive resource theories.

3.5.1. Checklist timings

Checklist completion times were recorded (Figure 28) from the onset of the list until confirmation of completion of the last item. If the checklist was interrupted by an ATC call to the crew's own aircraft, the time to deal with the call was excluded from the total duration of the checklist, allowing assessment of the effects of party line distraction without confounding ATC interruptions. Due to the large variance, the data were subjected to a log transformation before statistical analysis using a 2 (party line condition) x 3 (checklist) repeated-measures ANOVA. (There was one missing value for the party line condition of the descent checklist, as Crew 1 omitted the checklist on two occasions and did not confirm completion on the third.) The ANOVA revealed no main effects, but there was a significant interaction such that checklist duration for the descent checklist was greater in the party line than the no party line condition.

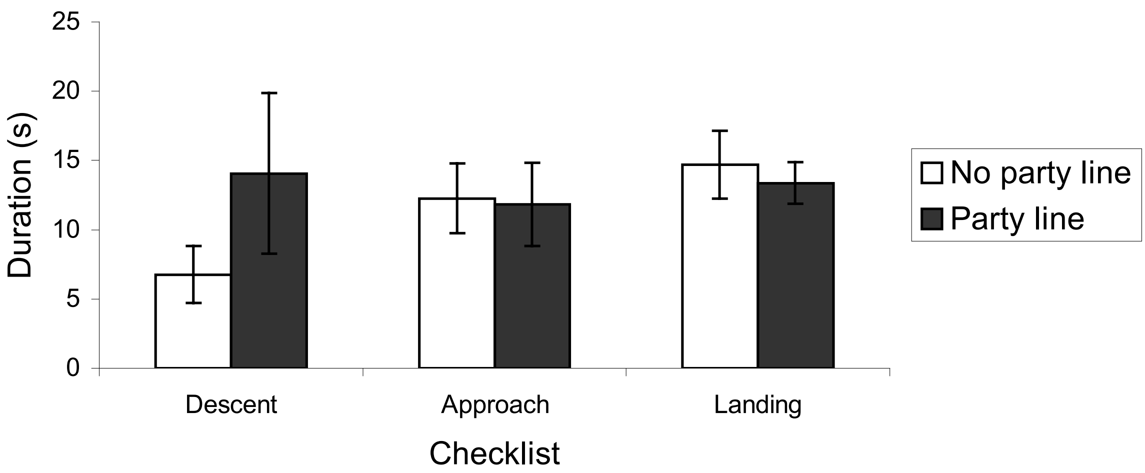


Figure 28: Duration of checklist(s) according to party line condition

3.6. Debriefing

Eight key questions aimed to facilitate discussion at the end of the day, eliciting pilots' opinions on the intrusiveness of background R/T in general, as well as their thoughts on the day's experiment specifically (see Appendix C). It was generally agreed that the party line condition caused greater distraction and increased workload relative to the condition with own R/T only. There was some disagreement as to whether this increase in distraction and workload affected the PF or the PNF to a greater extent. The party line was considered to be particularly distracting at busy times such as during the approach or at the top of descent, and generally at times when the workload was already high. It was not felt that flight safety was compromised in the party line conditions, but it was acknowledged that background R/T might increase the chance for mistakes. Pilots were able to give examples from their own flight experience of when the party line had been distracting, but also when it had been particularly useful in avoiding potentially dangerous situations. Hence pilots had reservations about the abolition of background R/T due to concern that situation awareness might be impaired.

4. DISCUSSION

4.1. Performance data

With respect to the influence of the party line condition, statistically significant effects were found only with respect to landing accuracy. However, this effect seems to be created by chance and reflects mainly the landing accuracy of the inexperienced F/O. In addition, during the landing phase, no background R/T was presented, and if there were an effect, it should be explained by the build-up during the whole approach. Examining the period shortly before touch down, the standard deviation of the roll and pitch angles do not confirm this hypothesis of lower performance in the presence of background R/T.

4.2. Physiological data

Heart rate, as expected, increased steadily for the PF throughout the duration of the flight, but there was little effect for the PNF. The data also reveal a slight trend for heart rate to be higher in the party line condition. Heart rate variability was lower as PF than PNF, indicating a higher workload for the PF. There was also a significant effect of phase, such that workload was lower in Flight Phase 1 (when the autopilot was on) than in the other four manual stages of flight.

Blink rate was significantly lower in the two PF conditions than in the two PNF conditions, especially during the intermediate stages of flight. This indicates that, as might be expected, the visual workload for the PF was higher than that of the PNF, and this workload increased after the autopilot was turned off.

$SP0_2$ was significantly lower in the PNF party line condition for the first flight phase, probably reflecting the preparation of the PNF for management of the R/T communications. For the other phases the trend was that the SPO_2 remained lower, indicating a higher workload condition. For the PF in the party line condition, there was a significant reduction in $SP0_2$ between the first and the final phases of flight. Almost no influence of the party line condition was apparent. These data parallel the heart rate results.

Skin conductance levels were significantly higher in the first two phases of flight than the final three phases. The data also indicate that, in the first two phases, skin conductance levels were higher for the two PF conditions compared to the PNF conditions. This higher activation corresponds with the First Officers' opinions that perceived time pressure and mental effort were much greater when flying as PF than as PNF, as expressed in the post-flight questionnaires. The SCL variability showed interesting results. Variability – an indicator of low arousal – was greater for the PNF party line condition than all other conditions during the intermediate stages of flight. This high SCL variability for the PNF party line condition then decreased significantly at the end of the flight, indicating a period of higher workload. This corresponds with the times taken to complete the landing checklist if compared to the no party line condition. The effect of lower workload in combination with background R/T is also mentioned in the debriefing: irrelevant RT allows the PNF to discard the messages until the advent of new meaningful information (e.g. start of a new ATC clearance).

The overall conclusion is that the physiological data indicate that, as expected, the workload for the PF was higher than for the PNF. The party line had some effect on the physiological responses. However, the effects were much less marked than was evident in the subjective data. In several cases the effects were even opposed to the subjective data.

4.3. Questionnaire results

The RSME and mental workload ratings indicate that the manual non-precision approach successfully imposed a relatively high workload. Crews reported moderate mental activity and perceived time pressure, and a slightly lower than average level of frustration. They judged their general success in accomplishing the flight task to be approximately 70%, and felt that R/T communication had contributed to a performance impairment of about 25% in the party line condition. All flights were perceived as being relatively safe, with a rating of 70%, assuming a baseline of fully automated flight.

A significant effect of party line was obtained for eight of the nine questionnaire items. Although largely irrelevant to the task in hand, the additional stream of background information increased pilots' selfappraisals of workload, mental activity, mental effort and perceived time pressure. In the party line condition, pilots were required concurrently to monitor and filter information in the irrelevant stream and so were not able to concentrate solely on the flying task. This created a multitasking situation in which more effort was required to co-ordinate the two behavioural streams, in comparison to each task performed separately. As expected, ratings of distraction and perceived performance impairment were also significantly elevated in the party line condition. The only questionnaire item not to be affected by noise condition was the perceived success in accomplishing tasks. Perhaps because pilots are used to the party line in everyday operations, they did not feel that task success was compromised by background R/T. However, although task success was perceived as comparable across flights, it was apparent from the ratings of the other measures that the crew had to work harder in the presence of party line to achieve this level of success.

There were also significant effects of flying role on a number of measures. Mental effort, mental and perceptual activity, mental workload and perceived time pressure were all rated as being greater for the PF than the PNF role, since demands are higher for the pilot in control of the aircraft. Conversely, the level of distraction from R/T communications was considered to be greater for the PNF than the PF, presumably because the PNF's main task is dealing with the R/T.

A near-significant interaction was obtained between party line condition and flying role for mental and perceptual activity. Although party line increased mental activity in both flying roles, this increase was significantly greater for the PNF than the PF. Since mental activity was already high for the PF, the introduction of background R/T had only a minimal effect. However, for the PNF, whose tasks are fewer, mental and perceptual activity increased significantly with the additional requirement to monitor and filter information in the background stream, compared to simply responding to ATC calls in the absence of the party line. There was also a near-significant interaction for perceived time pressure, such that the increased workload imposed by the presence of the party line augmented time pressure for the PNF, reaching the same level as that experienced by the PF.

It is worth noting the differences in subjective opinion according to pilot rank. Captains displayed significant effects of party line condition on ratings of mental effort, mental and perceptual activity, time pressure, frustration and success at accomplishing tasks, but First Officers did not. This stronger effect of noise condition for Captains might be accounted for by the fact that they were affected by the background R/T in both flying roles, whereas the First Officers generally reported that the party line had an effect only when they were acting as PNF. Accordingly, First Officers demonstrated significant interactions between party line and flying role, whereas Captains did not, on measures of mental workload and perceived time pressure, as well as trends towards interactions for mental effort, mental and perceptual activity, perceived performance impairment and distraction. These findings are consistent with comments made in the debriefing session, in which Captains claimed to monitor the party line in both flying roles, but First Officers attended to the R/T only as PNF. First officers were also more distracted by the party line than Captains, and were more likely to report that their performance was impaired by background R/T. These differences could perhaps be attributed to differences in flight experience. Main effects of flying role were obtained for First Officers but not Captains on two measures. First Officers rated time pressure and mental effort as being much greater when being PF rather than PNF.

4.4. Observational data

The observational data from the ATC calls indicated that slightly more calls were missed, queried or incorrectly read back in the party line than the no party line condition. However, the differences were small and there were insufficient samples to perform statistical analyses. Although actual errors were infrequent, the in-flight observer noted subjective observations of distraction. In particular, during crew briefings in the party line condition the PNF would often appear distracted, visibly trying to attend to both the PF and the background R/T. Call signs similar to their own (especially KLM 1024) appeared to capture attention. PNFs would sometimes pause momentarily before carrying out instructions (e.g., setting flaps), as if taking slightly longer to verify that the call was not for them.

Observations of checklists showed no apparent differences between the party line and no party line conditions. Omissions and repetitions of checklist items were perhaps more dependent upon *interruption* of the checklist (i.e., a call to the specific aircraft requiring action and read back) rather than mere *distraction* (i.e., the presence of background R/T). All the incidences of items missed or repeated resulted from an interrupted checklist. Interruption of the checklist did not always occur, and was equally likely to occur in the two party line conditions (four incidences of each). It would seem, therefore, that mere distraction caused by background sound was not enough to cause checklist errors; more important was the suspension and resumption of the list. The recent literature reviewed in the WP3 report supports the notion that interruption is a flight deck issue that merits much more attention.

Checklist timings showed an effect of party line for the descent checklist only. Although required to perform only two checks (terrain clearance and approach preparation), pilots often took some time to discuss these items and may thus have been more affected by the background R/T. In comparison, items on the other checklists involved quicker checks and less discussion, and therefore might have been less susceptible to distraction from the party line R/T. It was noted in the debriefing session that the checklists used were rather short; it is therefore possible that the effects of party line on checklist completion times would have been more apparent if longer checklists had been used.

5. IMPLICATIONS FOR THE COGNITIVE STREAMING MODEL

Before considering the implications of the results for the Cognitive Streaming model, it will be useful to reiterate the key features of the theory:

(i) Sound is subject to obligatory processing; that is, the sound is processed regardless of the will of the individual, so that even when the person is devoting attention to another task the sound will be registered and processed in various ways.

(ii) Sound that is subject to obligatory processing is organised perceptually. That is, the rules of auditory perceptual organisation — this organisation generally called 'streaming' — that are evident when a sound is processed consciously and deliberately are also at work when the sound is not attended. Hence the sound is partitioned in a way that reflects its 'object' properties, so that two distinct spatially-separate voices are represented as two distinct 'streams' of information in the brain. Reference to pre-attentive streaming (not requiring attention) should be qualified insofar as the rules of auditory perceptual organisation that apply to irrelevant sound are ones that apply to relevant sound when the person is listening passively, not when the person is deliberately trying to 'hear out' a sound, that is, when the listening is 'schema-driven'.

(iii) In addition to creating a record in the brain of the events (such as the words that are being spoken), pre-attentive obligatory processing and perceptual organisation yield a representation of the order of events. That is, a key element in sound processing is its organisation, and the major function of organisation is to produce a properly ordered sequence of sound. Moreover, this record is an enduring one — we may think of it being part of long-term memory. In this way, repeated sequences can be learned even when they are not being attended to.

(iv) The degree to which the processing described above in (i) to (iii) disrupts performance on an attended-to task can be understood in terms of the similarity of processing. That is, if the processing of the attended task emphasises short-term retention of order, then the degree of order produced by the unattended sound will determine the degree of disruption of the attended task. Thus cognitive streaming is distinctive inasmuch as it predicts that the disruption is due to 'similarity of process' (in this case the ordering — or 'seriation' — of items) not 'similarity of content' (that is, how similar the events are in the attended and unattended sequence).

(v) Another way of understanding the processing of order is to think in terms of 'transitional probabilities'. Series of events with low transitional probabilities are ones in which the likelihood of one event following another is low. For example, grammar and syntax in language help us predict what is coming next in a sentence. But the sequence of elements making up a call-sign or navigational co-ordinates, as examples, has no such predictability. Here, the brain has to do additional processing to retain the order in which the events were presented.

(vi) It follows from (iv) and (v) that, in the laboratory, activities that draw heavily upon short-term memory for order are particularly vulnerable to disruption by irrelevant sound. In the field, this should encompass activities that (a) involve dealing with novel information, (b) require short-term response to unpredictable events, and (c) call upon repetition of sequences (not just spoken sequences but sequences of actions also). Generally speaking, these factors will become more pronounced as the time pressure of events increases. Of course, the heavy burden on short-term order processing and increased time pressure typify 'high workload' conditions.

Broadly, therefore, we can expect tasks where the rate of information is high and the predictability of events is low to be particularly susceptible to the presence of party line.

5.1. Flight performance data

We may consider the results of the trial in relation to each of the dependent variables and evaluate their implications for cognitive streaming theory.

5.1.1. Periods of straight and level flight

- *Pitch angle standard deviations*: No significant effects

- *Roll angle standard deviation*: No significant effects
- *Flight director pitch bar*: No significant effects
- *Flight director roll bar*: No significant effects

Comment: Clearly, for straight and level flight the results are inconsistent with the theory, but this is a period where the burden of short-term memory is also low. Some of the performance changes may have been shaped by the attitude of the pilots to the trials and the fact that they were aware of the purpose of the experiment (that is, they were not 'blind' to the study's goal).

5.1.2. Landing period

- *Accuracy of final approach flight:* No deleterious effects of party line were found on final approach accuracy (pilots differed a great deal in their accuracy, as between Captains and First Officers, but these factors did not interact significantly).

- *Longitudinal touchdown precision*: There was a main effect of party line (such that there was a larger deviation in the party line conditions).

Comment: The result of the R/T influence on the landing performance is surprising since during the final approach almost no background R/T was present.

5.1.3. Overall comment on performance data

Only in the very late stages of flight do the data confirm the prediction, during the period of landing. Arguably, this is the very point at which moment-to-moment correction and responsiveness to the aircraft environment is at its most pressing; that is, the short-term memory load is greatest. However, this does not help explain why performance was better in party line conditions in earlier stages of flight. It is also difficult to explain the lack of significant effect for the pilot role during the landing phase. High workload is especially experienced by the PF, whilst no relevant ATC information is expected nor was presented.

5.2. Physiological Data

We may summarise the results from the five physiological measures as follows:

- *Heart rate*: The accumulating load of flying as the flight progressed was evident in the pilot flying in the heart rate measure. There was no significant effect of party line.

- *Heart rate variability*: No significant effect was found for the party line, although the values for the PNF in the party line condition seem to be consistently higher (meaning a lower workload with background R/T).

- *$SP0_2$*: This was lower for the pilot not flying and again there were no effects of party line.

- *Electrodermal activity*: The pattern of results is complex but shows generally arousal (or workload) to be higher early in the flight; the effects of the party line are rather inconsistent.

- *Eye blink rate*: This was lower for pilot flying conditions but there were no significant effects of party line.

- *EEG*: Here is some evidence that EEG levels may be higher for the PF (at least at the Oz electrode) for party line conditions in the final stages of flight.

5.2.1. Overall comment on physiological data

The cognitive streaming theory makes no precise predictions about physiological changes. However, it is predicted that the effect of the party line should interact with load. Inasmuch as the various physiological variables indicate workload, the results are somewhat disappointing with the exception of the EEG data (for which ironically one can be less certain that the precise pattern of changes reflects workload *per se*).

5.3. Questionnaire data

Again each measure is examined in turn.

- *RSME*: Pilots rated party line conditions as more effortful than no party line conditions.
- *Mental Effort and Perceptual Activity:* Pilots showed higher ratings in the party line conditions.
- *Perceived Time Pressure*: Pilots not flying showed higher ratings in the party line condition.
- *Perceived task success*: the party line did not show significant effects.
- *Mental Workload*: Pilots not flying showed higher ratings in the party line condition.
- *Frustration:* The party line led to significantly higher frustration levels.
- *Distraction*: The party line produced markedly higher ratings than no party line (this time the effect of flight role was marginal).
- *Performance Impairment*: Interestingly in the light of the performance data, in both roles the aircrew thought that the party line had a deleterious effect on their performance.
- *Perceived Flight Safety*: This was reduced in party line, particularly for the pilot not flying.

5.3.1. Overall comment on questionnaire data

Comments and qualifications about the subjective ratings have already been made. Cognitive streaming theory does not make predictions about the effect on subjective ratings; it could be argued that sound will intrude and have its effect without necessarily leading to awareness of its effects. Certainly, the results are in line with increasing the awareness of disruption. However, the fact that the effects are most clearly manifest in pilots not flying is notable.

One of the PNF's major roles is dealing with R/T messages. This distinction between the PF and PNF roles may be crucial. However, if the processing of irrelevant sound is obligatory, as demonstrated in a great number of previous experiments, it is difficult to explain why the PNF should be selectively affected by the party line. One possibility is that the pilots flying were too busy to formulate views during the flight and the pilot not flying had time to reflect on how the sound was impairing performance. Perhaps most interesting of all is that the pilots flying thought that there should be a material effect of party line sound on flight performance and safety, but the effects on performance are actually rather inconsistent and small.

6. CONCLUSIONS

The systematic subjective reports are unequivocal in showing that the aircrew reported that the effects of party line were negative. Ratings related to workload and distraction were higher in the presence of irrelevant R/T messages. Although the cognitive streaming model does not make firm predictions about subjective response to sound, these results are consistent with the notion that an irrelevant stream of information is difficult to ignore.

The subjective effects were not reflected in changes in flight performance. Only in the very last stages of the flight were some negative effects of party line made manifest; elsewhere, in the earlier stages of flight, the party line seems to have had a beneficial effect. One way of understanding the flight data is to suppose that the effects of workload and distraction were partly cumulative, so that as the demands of the task were greatest — when the pilot had to land the aircraft — the predicted effects were found. Alternatively, the nature of the tasks performed in the various phases of flight may account for differences in the effects of the party line.

Again, the physiological data were mixed. Some of the physiological measures were sensitive to workload differences, as expected. The effects of the party line may have been much smaller than those of workload generally, and thus have been ‘washed-out’ in the variance.

The results of this study are rather different from those in previous laboratory-based research. The typical pattern of results in the laboratory is that performance is substantially degraded by irrelevant sound, the level of degradation often reaching 30%. The participants, however, may have no awareness of this effect. In the present study, in contrast, experienced aircrew reported a very clear subjective effect of the party line, but some of the objective measures indicated performance *improvement*.

A relatively small sample size was inevitably used in this study, but the data suggest that statistical power was adequate: many significant effects were achieved. The results therefore cannot easily be dismissed. Differences between the laboratory and the current 'real-life' conditions that might have contributed to the discrepancy include:

- In laboratory studies, well-controlled but artificial tasks are typically used, whereas the flying task is more complex and less tightly controlled.
- In laboratory studies, the participants are generally relatively unfamiliar with the task, whereas aircrew are highly practised in flight skills.
- In addition to differences in expertise, we might expect differences in level of motivation; aircrew naturally take great pride in their professional skills, whereas participants in laboratory studies have little 'ego involvement' in their tasks.
- Predictability for pilots of the events in real-life tasks is high, but for non-pilots in the laboratory it is low. The party line may have increased the realism in the present study and hence the participants' vigilance, even though their practical experience tells them that party line is disruptive.
- Laboratory and real-life tasks often call upon similar processes such as keeping track of order, but experienced aircrew may develop methods of scheduling sub-tasks that are not available to the laboratory participant.
- It tends to be relatively easy to devise performance measures for laboratory tasks, whereas performance in flight simulation is much more difficult to quantify. Interestingly, a simple measure of checklist completion time showed the expected decrement in the presence of the party line.

Whatever the source of the findings reported here, it certainly cannot be concluded that the irrelevant sound effect is simply a 'laboratory' phenomenon. Substantial decrement in performance has been reported, for example, in open-plan office environments (Banbury & Berry, 1997, 1998).

The results obtained here contribute to the growing body of evidence concerning the nature of cognitive streaming. As the theory develops, it may be possible to account more definitively for the present results. Meanwhile, theoretical development will be assisted by the wide range of measures provided by this CARE trial.

7. REFERENCES

Banbury, S., & Berry, D.C. (1998). The disruption of office-related tasks by speech and office noise*. British Journal of Psychology*, 89, 285–295.

Banbury, S., & Berry, D.C. (1997). Habituation and dishabituation to speech and office noise. *Journal of Experimental Psychology: Applied*, 3(3), 1–16.

Farmer, E., Chapman, V., Brownson, A., Thompson, D., & Jones, D. (2003). *CARE Innovative Action Cognitive Streaming Project: Report on Work Package 3: Prevention of Performance Impairment*. Report to Eurocontrol, Ref CARE-IA-CS-QIQ-WP3-D3-01-1.0.

Hoogeboom, P.J. (2003) Off-line synchronization of measurements based on a common pseudorandom binary signal*. Behavior Research Methods, Instruments & Computers,* 35, 384–390.

Jones, D. & Farmer, E. (2001)*. Applying the Cognitive Streaming Theory to Air Traffic Management: A Preliminary Study.* Report prepared for Eurocontrol, ref QinetiQ/CHS/P&D/CR010575/1.0.

Jones, D.M., Macken, W.J., & Mosdell, N. (1997). The role of habituation in the disruption of recall performance by irrelevant sound. *British Journal of Psychology*, *88*, 549-564.

Jones, D., Parmentier, F., Houghton, R., Pope, D. and Farmer, E. (2002). *Possible Usefulness of Irrelevant Sound.* Report to Eurocontrol, Ref CARE-IA-CS-CFU-WP1-D1-01-1.0.

Lykken, D.T., Rose, R., Luther, B. & Maley, M. (1966).Correcting psychological measures for individual differences in range. *Psychological Bulletin*, 66, 481–484.

Macken, W.J., Tremblay, S., Houghton, R.J., Nicholls, A.P., & Jones, D.M. (2003). Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory**.** *Journal of Experimental Psychology: Human Perception and Performance*, 29(1), 43–51.

Pope, D., Houghton, R.J., Jones, D.M. & Parmentier, F (2003). *Report on Work Package 2: Range of Tasks Affected by Irrelevant Sound*. Report to Eurocontrol, Ref: CARE-IA-CS-CFU-WP2-D2-02-1.0.

Prinzel, L.J., Freeman, F.G., Scerbo, M.W. Mikulka P.J. & Pope, A.T (2003) Effects of a psychophysiological system for adaptive automation on performance, workload and the event-related potential P300 component, *Human Factors*, 45(4), 601–613.

Wickens, C.D. (1992). Engineering Psychology and Human Performance. Columbus: Charles E. Merrill.

Yagi, A., Kazai, K., & Takeda, Y. (2000). Spatial and temporal variations in eye-fixation-related potentials. In A Yagi (Ed.), *(Special Issue: Brain Images and Cognitive Functions) Japanese Psychological Research*, 42(1). Blackwell Publishers (ISSN 0021-5368).

APPENDIX A: BRIEFING GUIDE FOR PARTICIPANTS

Cognitive Streaming

(Care Innovative Action)

Briefing Guide

Generic Research Aircraft Cockpit Environment
NLR, Amsterdam

February 2004

Table of Contents

1. Introduction

Welcome to the National Aerospace Laboratory NLR, based in Amsterdam. NLR is an aerospace research establishment, studying research topics ranging from aerodrome safety to space applications. For the present experiment NLR has invited you and a colleague pilot to participate a study that focuses on the development of useful assessment methods in Human Factors. You will be flying in the GRACE (Generic Research Aircraft Cockpit Environment), a fixed-base research simulator, programmed to simulate the Boeing 747-400. The flights are part of the CARE Cognitive Streaming project. This is a EUROCONTROL-funded project aimed at developing innovative technologies in aviation. This briefing guide contains information about the project, simulator and equipment that will be used, and on operational procedures that you will be following during the flights.

Important comments and specific requests from the researchers are given in bold and italics, and are summarised in Section 6.5.

Please answer the questions in <appendix A> and <appendix B>, and bring the forms with you to the NLR.

2. Background and aims of Cognitive Streaming

Cognitive Streaming is part of the Eurocontrol CARE project that aims to demonstrate how a new theoretical approach to understanding human information processing can be applied to the civil airliner cockpit. The cognitive streaming theory explains how human speech, even when unattended to, can interfere with task performance that requires processing by the human brain. This is mainly due to the fact that human speech is processed phonetically to determine whether it contains relevant information. In this way, ‘streams’ of phonetically coded information are created. These streams of information can easily be confused with streams containing other information, created by current task performance. For example, streams of ‘to be ignored’ radio-telecommunication (R/T) can interfere with streams of information of ongoing tasks like executing checklists. That is why the present study will test the effects of ‘partyline’ on aircraft operations.

More specifically, we will measure the effects of partyline disruption on crew performance and information processing. This will be achieved by linking aircraft performance with information derived from questionnaires and physiological data collected during non-precision approaches at Schiphol airport. Six landings will be performed with partyline R/T present and another six landings will be performed without the partyline.

The next chapter will give an overview of the cockpit simulator. Chapter 4 will discuss in detail how the measurements are performed. Finally, Chapter 5 provides the research operational procedures essential logistical data for the experiment day.

3. The Flight Simulator GRACE

3.1. General Layout

The Generic Research Aircraft Cockpit Environment, GRACE, is a generic flight simulator, representing a modern large airliner. The current configuration is Boeing 747, with Boeing 747-400 enhanced displays and a simulated Boeing 747-400 Flight Management System.

The main simulator components of the GRACE are:
- generic cockpit
- control loading system
- visual system
- sound system
- control desk

3.2 Cockpit Displays, general

The cockpit features in total six displays, which are equivalent to existing Boeing 747-400 cockpit displays. The Navigation Display is controlled with the Display Control Panel (DCP) which is located next to the Autopilot Mode Control Panel. The layout of the DCP with its functions is shown below.

3.3. Flight Management System

The Flight Management System is a simulation of a Boeing 747-400 FMS, which is connected to two Control & Display Units (CDU) which are fitted in GRACE. When flying with the autopilot and FMS engaged the aircraft will be guided along the predefined route. Through the cockpit CDUs and displays the pilots will monitor flight progress and implement route changes.

3.4 Autopilot/Mode Control panel

The autopilot of GRACE is similar to the Boeing 747-400 autopilot, and is controlled through the Mode Control Panel as shown below.

With amongst others from left to right:
- Speed control (AT and speed select)
- LNAV
- VNAV
- Lateral control (heading hold/select)
- Altitude control (vertical speed, altitude hold/select)
- Approach control (Localiser and approach)
- Autopilot engage/disengage

4. Measurements

4.1 Overview

The following data will be collected during the experiment:

- Video recordings of cockpit activity, including communication (between crewmembers & ATC transmissions)
- Digital recordings of brain activity and eye movements (to determine whether auditory information is processed)
- Non-invasive measurements of oxyhaemoglobin saturation (SPO2), heart rate and respiration (for mental workload determination)
- Simulator data logging for the determination of aircraft performance
- Observations and questionnaires asking for your opinion about events that happened during the flight.
- Self-ratings by pilots in a debriefing session

The measures will be taken from both crewmembers.

All results of the experiments will be dealt with anonymously. No personal or company names will be related to any of the outcomes of the experiments.

4.2 Video recordings

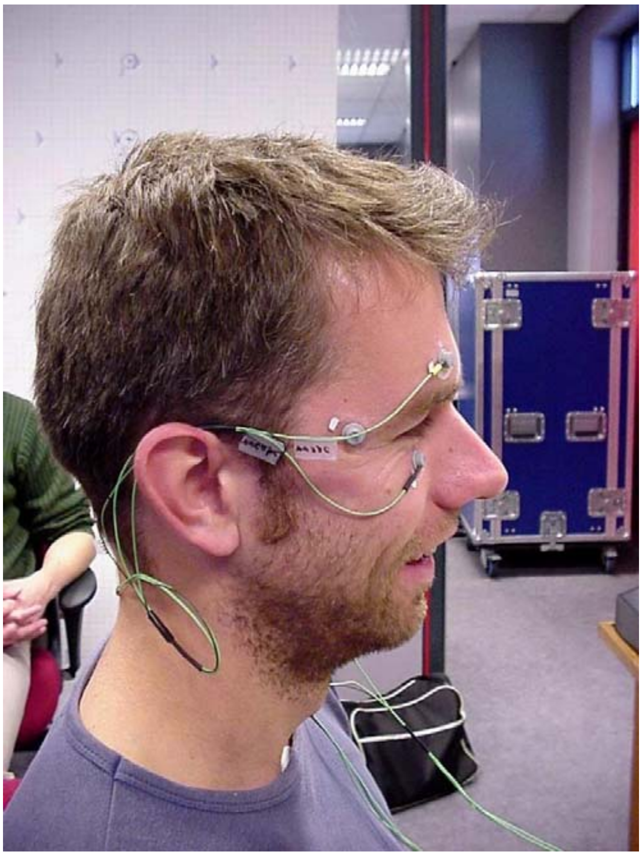
The video recordings will be used as a check of the events that occurred in the cockpit, like the exchange of information or responses to calls from ATC. The main focus will be on auditory information.

We kindly ask you to adhere to the crew procedures as you would have done under operational circumstances. Always follow Standard Operating Procedures (SOPs), even if the events that occur may seem like simulator errors.

4.3. Determining party line information processing through brain activity

Electroencephalography (EEG) measurement equipment will enable us to tell whether and in what extent the brain processes sounds during the partyline. The EEG equipment consists of 3 electrodes that will be placed on your head, 4 around your eyes, and 2 on your ears as a reference. We will use collodion to place the electrodes on your head. Collodion is a gel that hardens, but can also be easily removed with special remover.

We kindly ask you not to pull the electrode cables since they will be securely placed on your head. Could you please inform the in-flight observer in case you feel any discomfort or irritation as a result of electrode placement.



Picture of eye movement electrodes

4.4 Registration of changes in workload via heart rate

Using physiological measurements an objective measure of mental effort (mental workload) can be obtained. This is why heart rate but also respiration and galvanic skin response will be recorded. Heart rate is measured by placing electrodes on your chest, and respiration is measured through a stretchable cable placed around your chest. We would like to emphasise that these measurements are not for any medical or bio-medical purposes.

Please do not perform any heavy physical activities on the day before the experiment.

Please wear a T-shirt under your clothes so that we can easily place the heart rate electrodes.

4.5 Simulator logging of aircraft performance

The relevant parameters of the simulated aircraft will be recorded to enable a reconstruction of the flightpath, and events that may have occurred.

We kindly ask you to fly as you would have done under normal operational circumstances.

4.6 Questionnaires

We will present you with a number of questionnaires before, during and after the simulations. The main purpose is to get feedback from you. The following questionnaires will be used:

- **'Start-of-day' questionnaire.** The purpose of this questionnaire is to obtain biographical information, information about flight experience, etc.
- **Post-flight questionnaires.** The purpose of these questionnaires is to obtain your feedback on performance related to the specific flight. The questions will be presented to you in between the landings performed in GRACE.
- **'End-of-day' questionnaire.** The purpose of this questionnaire is to obtain your general opinion about the level of disruption caused by the partyline but also your level of fatigue.

5. Crew Operating Procedures

5.1 General description of operations

The relevant approach and landing procedures are given during the GRACE training at NLR in Amsterdam.

The procedures for the operation of the simulated aircraft (Boeing 747) differ from the SOPs normally used for airline operations. Most differences spring from the fact that GRACE is not a training simulator, but a research facility. The cockpit contains no systems controls (all hydraulic, pneumatic, electrical and environmental systems on board the GRACE will always work), so many normal and most abnormal checklist items pertaining to systems can be skipped.

The Cognitive Streaming checklists have been included in this briefing guide as Appendix A. The memory items in the emergency checklist need not be learned by heart, for any emergency procedure requiring memory items will be included in the practice flights, if applicable. Below follows a summary of important operating information, presented in chronological sequence. The relevant Company Crew Co-ordination (CCPs) are included in this summary.

CCPs that are generic to most aircraft in the fleet will have to be followed during the high- workload stages of the flight to ensure proper and smooth aircraft handling. During the training you will be given a summary of important operating information.

We kindly ask you to strictly adhere to the crew procedures as you would in a normal non-experimental flight.

5.2 Beginning of simulation

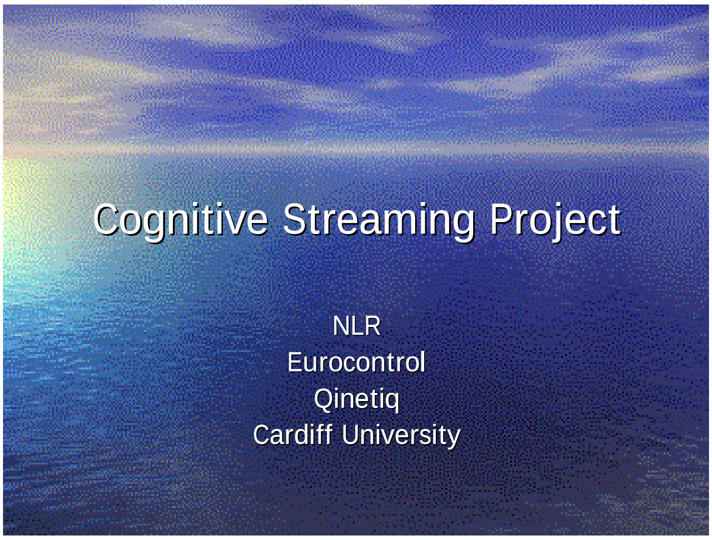
5.2.1 Descent towards Schiphol

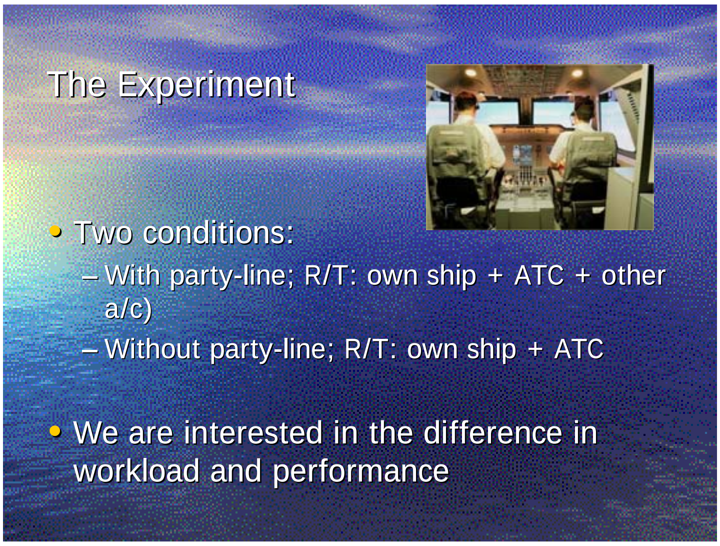
INITIAL POSITION: FL140, 30 KIAS at SULUT
REDFA ARRIVAL
SULUT
SUGOL
30nm SPL
15nm SPL
RADAR VECTORS to INTERCEPT NDB DME approach RWY 18C

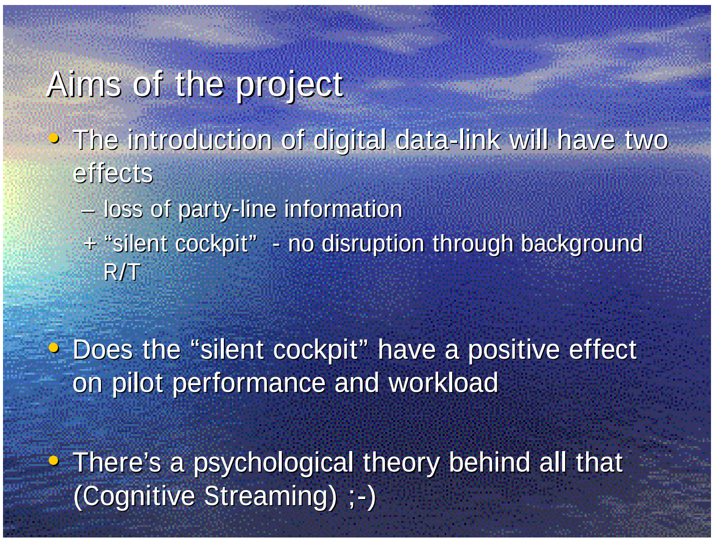
5.2.2 Landing

Automatic GPWS call-outs are operational: 500-100 50, 40, 30, 20, 10 and will aid in estimating when to start the flare. The quality of the touchdown and roll-out are a part of the experiment.

Slides used for pilot briefing on Day 1

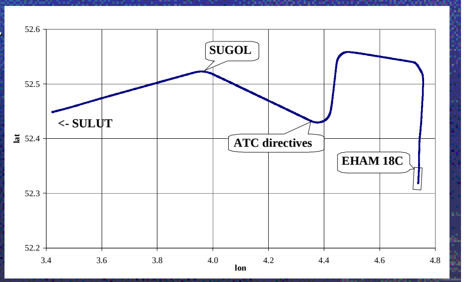






Flight: Approach EHAM 18C via REDFA

- Initial position: FL140, 300kias
- Fly according to flight plan (analysis)
- normal R/T
- A/P disc at SUGOL
- Radar vectors
- LDG 18C



Measuring performance and workload

- The flight simulator GRACE: Six landings will be performed at Schiphol airport.
  - With and without party line.
- All data will be dealt with anonymously.
- Please adhere to normal operational procedures.
  - SOP/CRM
  - checklists
  - R/T
- Simulator data will be recorded.
  - e.g., flight path, events, speed, altitude.

Measuring performance and workload

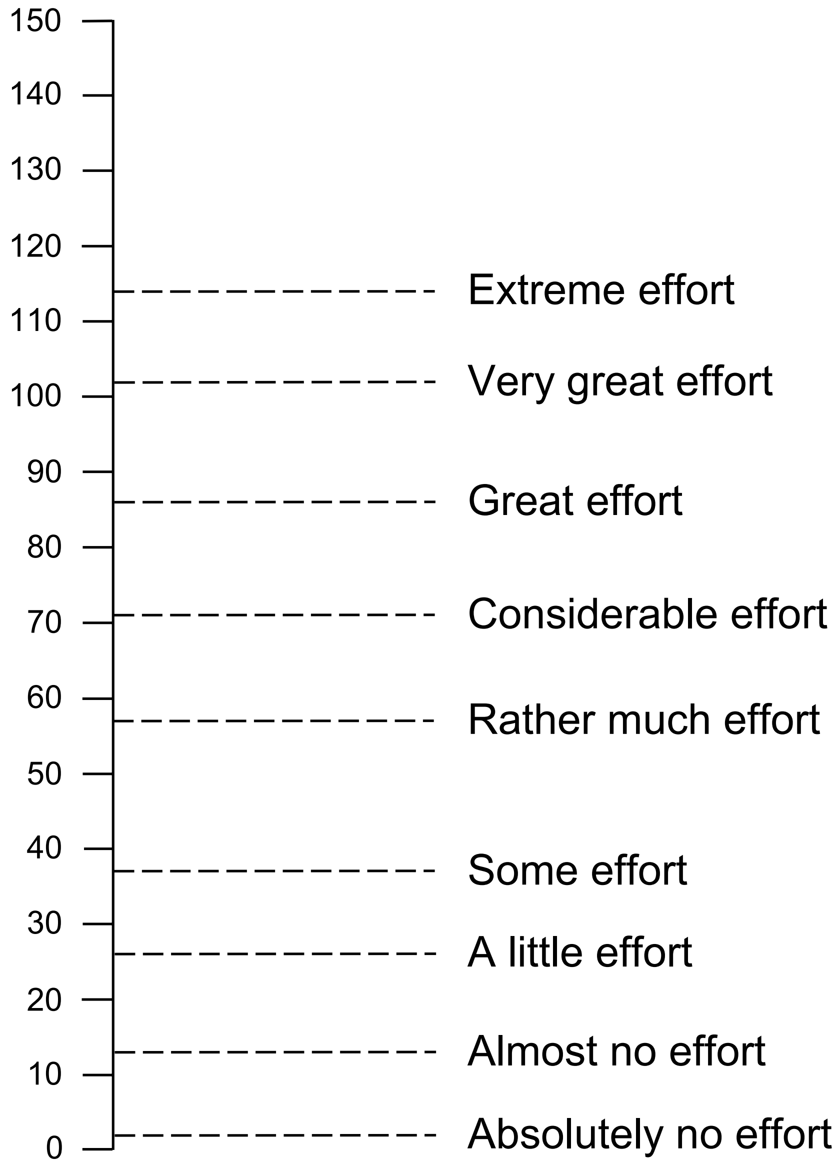
- Observations/ video recording of cockpit events.
  - Please speak in English.
- Digital recordings of brain activity, eye movements and other physiological indicators.
  - Brain Activity
  - Eye Movements
  - Heart Rate
  - respiration and movement
  - and other physiological data

Electrodes will be placed around your head, eyes and ears (assessment of information processing).

APPENDIX B: RATING SCALES

Rating Scale Mental Effort (RSME)

Please indicate, by placing a mark on the vertical line below, how much effort you had to invest in order to execute the task (that you have just been working on).



CARE Cognitive Streaming

crew	position	Run ident	date

1. How much mental and perceptual activity was required (thinking, deciding, calculating, remembering)?

0 50 100

2. How much time pressure did you feel due to the rate or pace at which the task elements occurred? For example assess whether the pace was slow and leisurely or rapid and frantic.

0 50 100

3. How successful do you think you were in accomplishing the tasks you were required to perform during this exercise?

0 50 100

4. How hard did you have to work (mentally) to accomplish your level of performance?

0 50 100

5. How insecure, discouraged, irritated, stressed and annoyed (versus secure, gratified, content, relaxed and complacent) did you feel during the task?

0 50 100

6. To what extent did R/T communication distract you while you were performing your task?

0 50 100

7. To what extent do you think R/T communication impaired your performance?

0 50 100

8. To what extent did you feel that the flight you performed was safe?

0 50 100

APPENDIX C: DEBRIEFING

To what extent did background R/T add to your workload? Was the situation different for PF and PNF?

Although all pilots agreed that the party line increased workload, there was a lack of consensus as to whether this affected the PF or the PNF to a greater extent. Crew 1 believed that monitoring the background R/T was more the task of the PNF than the PF, and hence suggested that the presence of the party line added to each pilot's workload 30% and 10% respectively. Crew 2 indicated that background R/T increases workload by about 20% compared to the silent cockpit condition, and that this increase is the same for both crew members. Crews 3 and 4 were divided in opinion, with the First Officers in each more inclined to say that the party line added to the workload of the PNF, whilst the Captains believed that R/T had the greatest impact for the PF. This difference in opinion reflected differences in perceived roles of the PF and PNF in the cockpit. First officers saw the tasks as being divided, with the PNF dealing with all the R/T and the PF being responsible for the actual flying of the aircraft. On the other hand, Captains saw the monitoring of the R/T as an additional task for the PF, as it is the PF who is in charge of the aircraft and so it should be his role to know what settings are required.

- "I especially noticed it when I was PF, because the standard operating procedures tell you to give him the orders to set things on the panel and so you have to tell him what to do…so you have to listen to what ATC say." (Captain, Crew 3).
- "As a PNF the background R/T adds quite a lot of workload, but as PF I didn't really notice much difference between with and without background R/T." (F/O, Crew 3).
- "You both have to listen to the R/T call which is meant for you, and in principle you both should be able to read it back to the controller. When you are PF, it is an extra increase in workload and then of course it is a higher workload when you have party line as well." (Captain, crew 4).

Did you feel less distracted in the condition without the R/T background communication? Was the situation different for PF and PNF?

It was generally agreed that the condition without background R/T was less distracting, although the more experienced pilots were keen to note that the party line was something they were used to, and did not feel that it degraded their performance.

- "I am used to the background R/T in day to day practice…I do like the background R/T because I know in that case where the other aircraft are. It is indeed distracting when you want to switch something or do something, you have to listen to it…so something of your brain is working on that part…but it's not distracting in such a way that I don't feel comfortable about it." (Captain, Crew 4).

Again, there was some difference of opinion as to whether the presence of the party line had a greater effect on the tasks of the PF or PNF:

- "I was a little less distracted, especially as PF." (Captain, crew 2).
- "Yes, there is more distraction (with party line), especially as PNF." (F/O, crew 3).
- "Yes it was less distracting without the R/T in the background. But this I felt only while being PF because you try to listen to R/T as well. As PNF it did not matter to me, because you're only scanning the PF." (Captain, crew 3).
- "Yes, PNF does the R/T, so during monitoring he has to continuously process all R/T information. For PF interference is less." (F/O, crew 4).
- "The situation is worse for the PF because his workload is higher than the PNF." (Captain, crew 4).

The First Officer of crew 1 believed that background R/T was only distracting at points of high workload; when workload was low, the party line could actually be beneficial by keeping the PNF alert. The Captain however, disagreed, saying that he always found the background R/T distracting. One factor that all crews referred to was that similar call signs could be particularly distracting:

- " If you hear KLM and it's not your call, you are still distracted for a couple of seconds." (F/O, Crew 3).

Some pilots mentioned filtering out calls in the party line based on the early part of the call sign; of course, the more similar the call sign, the longer attention is captured before the call can be dismissed.

- "I only listen to the first part but it takes a little time…every time you have to listen to it and you are distracted from your other task. Today you used KLM 1024, but when the '1' came up I knew OK, that's a European flight, and we're 204. So when I heard '1', I knew it wasn't me." (F/O, crew 3).
- "The controller starts talking, gives the call sign, then you think 'oh, that's not me' and let it go. You don't really think about it but that's the way it works. Whenever you hear your own call sign then you think 'what's this?'" (Captain, crew 2).

However, some pilots said that they did not filter out the messages, and instead chose to listen to them so that they knew what other aircraft were doing.

- "You tend to listen because you want to know what happens." (Captain, crew 3).
- "I noticed there were a lot of calls, 2014 and things like that but that's not the problem. It's that you have to process all the information because first you want to know who it is for, and even if you confirm to yourself its not your flight, then you still continue listening to the message…because you like to know, maybe a message is also a landing clearance, a clear for the same runway." (F/O, crew 4).
- "It is distracting as you have to process it, it takes some time away from the things you're doing in the cockpit…For me it is not possible to disregard it totally, I also listen to the entire R/T call." (F/O, crew 4).
- "During some parts of flight you don't filter it all out, because you want to know what is being said…but I have two ears and one ear I can split from the other one, so I can hear calls for the other aircraft. That took me four or five years."(Captain, crew 4).

Although pilots noted the benefits of the party line for situation awareness, they also acknowledged that the information they hear is not retained for very long.

- "It's nice to have party line, but when you're busy you're not paying attention to other calls anyway…if you ask someone ten seconds later what clearance another aircraft just received, they won't know. You remember for a few seconds, then it will be gone." (Crew 3).
- "I hear the clearances for the other aircraft but I cannot read them all back, not for the ones that aren't my flight." (F/O, crew 4).

Pilots were in agreement that the party line condition was more distracting because it involved constant monitoring of the background R/T, where as in the silent cockpit condition they knew that every call received was relevant.

- "It is more relaxing because you don't have to focus all the time." (Crew 3).
- "I was more distracted whether I was PF or PNF with the party line, but it is not realistic to compare it to a flight with nothing, because then you hear nothing at all and when you hear something you know its for you". (F/O, crew 4).

Did you find that the party line and background speech might have been more distracting at some points than others? If so, when?

Crew 2 reported a similar level of distraction throughout the duration of each trial, as the scenarios used (i.e., approaching the airport and landing) involved the busiest parts of flight.

- "Normally in real life in the cruise flight, it's not very distracting and it's quieter anyway. Whenever you get close to an airport it's more distracting." (F/O, crew 2).

More specifically, the start of descent was repeatedly mentioned as a point when background R/T could be particularly distracting.

- “I started the descent late and I really noticed that was because I was listening to R/T…we were already past it by almost a mile.” (Captain, crew 3).
- “With the party line basically on the last part when you finish your approach and start your descent, that’s when you actually have to fly manually and it becomes more distracting.” (F/O, crew 3).
- “With party line R/T you are distracted during some critical points…so maybe setting altitude, during the top of descent and gear down, can sometimes be difficult.” (F/O, crew 4).
- Crews also referred to the approach as a time when background R/T can be a problem.
- “It is distracting when selecting flaps and decelerating, just before and during the approach.” (F/O, crew 3).
- “Just before approach is distracting. I think it really becomes a real subject if the workload is already high. If you don’t have much to do, you have all the time and you have a lot of space in your brains to react, but just before intercepting the approach when doing checklists and stuff, then it becomes a subject.” (Captain, crew 3).

In general, it could be said that the background R/T is more distracting when workload is high. The scenarios used in the experiment exploited this by increasing workload through the use of non-precision approaches and the disconnection of the auto pilot.

- “I think it also depends on if you are making a precision approach or a non precision approach…on a non precision approach there are a lot more checks to do to see if you’re on the right descent profile…you have to do a lot of things at the same time, so that really becomes distracting if ATC makes a call.”(Crew 3).
- “Without the auto pilot, the ATC can get a bit too distracting sometimes.” (Crew 3).

Did you feel that the flight’s safety was different in the conditions with/without background R/T?

Crew 3 again mentioned the flight in which they started their descent late, which the Captain attributed to R/T distraction.

- “We were one mile past our descent point, so 300 ft high and still flying level. We had to start our descent and that’s getting dangerous…I noticed that I was paying more attention to ATC than to my descent profile.” (Captain, crew 3).

The first crew noted that in the party line conditions of the early flights, their performance was less precise and that sometimes they were flying below speed. In general though, pilots did not feel that safety was compromised.

- “It was easier to concentrate and focus without background R/T, but background R/T did not impair flight safety in my opinion.” (F/O, crew 2).
- “There was a higher workload but it did not feel unsafe” (F/O crew 4).

However, it was acknowledged that mistakes could perhaps occur in the presence of a lot of background R/T.

- “There is a possibility when there is a lot of party line that you could miss a call… during critical moments like the passing of a flight level or a NDB clearance. It’s not consistent during the entire flight when you have party line that your safety’s impaired. If it happens I think it’s during the critical moments.” (F/O, crew 4).
- “With background R/T the chance of accidents is greater because mistakes in heading, speed and level can be made. Without background R/T you know when you hear a controller it’s a call for you.” (Captain, crew 3).

Flight experience may be an important factor: First officers were more likely to say that the party line had a negative effect on their performance, whilst Captains tended to see it just as an inconvenience.

- “For me, to some extent as I’m not used to flying with so much R/T.” (F/O, crew 4)

- "It doesn't influence me that much, maybe a little, but that's only because at the moment I want to give a command to lower the gear or select some flaps for instance, and there's an R/T call, then I have to wait for him to finish the R/T call before I can give the commands." (Captain, crew 4).

On the other hand, pilots did also note the potential benefits of background R/T to flight safety in terms of situation awareness.

- "I don't think safety really improves without party line. You get less distracted but on the other hand I think it's very important to have a good mental picture of what's going on around you and the R/T of other aircraft helps." (Captain, crew 2).

Can you think of any examples from your experience where party line and background R/T distracted you (and caused a potentially unsafe situation)?

Crew 4 believed that background R/T is only really a problem when workload is very high, for example, when situations have involved engine fires etc (in training). In normal situations however, they had no experience of background R/T contributing to a potentially unsafe situation.

As discussed earlier, the fact that airlines use a lot of similar call signs means that pilots have to work a lot harder monitoring the R/T for their call. Although they did not feel that this creates an unsafe situation, it can increase workload when approaching busy airports
.

- "I had annoying background R/T when I had to do a single pilot approach at Schiphol…You have a KLM flight number…at Schiphol I think about 80% of all the call signs are KLM, and you hear about six of them a minute, so if you have to do everything on your own and you're so distracted…I think it takes about 60% of your total brain capacity to just get the R/T calls." (F/O, crew 3).

Crew 2 mentioned that missed calls were quite a common occurrence, but it is not often a problem as pilots will cross check with each other, or ask ATC to clarify if they are still unsure.

- "It happens quite a lot actually. In a normal day to day operation, lots of calls get missed and read back wrongly." (Captain, crew 2).
- "I read back for another aircraft once but ATC noticed it so it wasn't a problem, but if they don't notice it then it's a problem." (Captain, crew 3).

Crews 1 and 2 both mentioned that potentially unsafe situations can be created when the frequency is constantly busy. For example, if a landing clearance is needed, then the other R/T can interfere with your transmission.

Can you think of any examples from your experience where party line information helped you to detect important information (and resolve a potentially unsafe situation)?

All crews agreed that the party line is useful as it increases situation awareness. This can be particularly helpful at bad or busy airports as it makes it possible to anticipate potentially unsafe situations sooner, therefore giving more time to react. Pilots were able to give a few examples:

- "Back in America I was clear for the right hand base but the tower didn't have me in sight, and another aircraft was clear for the left hand base but the tower didn't have him in sight either. I turned my base, flew it and I was straight on with the other aircraft but I was a little expecting it because I heard 'left hand base I do not have you in sight', and I thought I should pay special attention to that. The plane was really close, but I was able to avoid it…I don't know if I would have seen it at the right time if I did not have the party line." (F/O, crew 3).
- "I was established on a final for a runway at Cardiff I think. ATC let other traffic pass in front of us and they did not tell us exactly what was going on. Just by listening to the other guy we knew what was going on, and then we had time to react on this to prevent an actual near miss." (Captain, crew 2).
- "On one occasion the party line helped prevent a near miss with crossing traffic, a helicopter, when we were established on the final approach." (Captain, crew 3).

- “During PRM (position radar monitoring) approaches, that is an approach on two runways at the same time, with two aircraft beside each other within 405 metres of each other. Then it’s extremely useful that you have his R/T as well as yours, so that you are aware of his position.” (Captain, crew 4).

As well as issues of flight safety, it was also noted that the party line could be useful for in flight planning and flying economically.

- “It can be really useful to fly economically, because if you hear that, lets say for instance, one aircraft is number 2 for the approach and there’s nobody else, you can reckon that you will be number 3. If it is busy you can already reduce your speed.” (F/O, crew 3).
- “You can anticipate by helping ATC, because if you hear them say something to another aircraft and you can see them move on the radar, you can anticipate OK, he’s coming in, I’m still flying fast… I’m gaining on the one in front of me, so I’m reducing, so then you also reduce the workload for the ATC, because if you’re still fast he has to tell you slow down.” (Captain, crew 3).

What did you like best about the experiment? Are there things that you would like us to change?

It was generally agreed that the experiment was interesting and well set up, and that the information provided beforehand was useful. Crew 4 suggested that for less experienced pilots, more information might be needed in the briefing guide about the flight instruments. It was felt that the simulator was good and realistic enough for the purposes of the experiment, although motion would have been helpful as pilots reported a tendency to overcorrect when judging the approach. The workload was considered to be acceptable, and some suggested that additional workload and perhaps abnormal situations could have been used in order to bring out the effects of background R/T further. It was also suggested that different scenarios could be used, as the workload was reduced considerably in the later flights when pilots became familiar with the approach that they were performing.

The amount of R/T was considered fairly representative of a normal, non-rush hour approach into Schiphol. It was suggested that the two conditions (with and without party line) were somewhat unrealistic to compare, as the no party line condition does not require any monitoring of the R/T at all.

- “Maybe it’s better to have one with a lot of background R/T and one with less or at least some other aircraft, as right now when you hear something you know it’s for you.” (F/O, crew 4).
- “You don’t have to figure out if you’ve got the right call sign or not, every call which is made is for you.” (F/O, crew 2).

It was also frequently mentioned that the R/T in normal flights can be very poor quality, whilst the ‘KLM 204’ call in this experiment was very clear. Pilots said that trying to pick out important words from the background R/T can be a more effortful process in a real flight, and therefore leads to more missed calls and incorrect read back.

- “I would like to emphasise that R/T in live aircraft is worse to understand and therefore consumes more energy and concentration.” (F/O, crew 2).

Are there any other comments that you would like to make?

Pilots had mixed views as to whether they would prefer to have the party line or data link technology. Although the Captain of crew 1 welcomed the abolition of background R/T, the other pilots had more reservations about the loss of party line information and considered the possibility of combining R/T and data link.

- “I would prefer data link together with party line so clearances for me should appear on a screen, but how is that possible? How can you create a situation so that I hear party line R/T and I get my own clearances on a screen? That’s not feasible. Having all clearances on a screen is too cluttering, I’d rather have it in my ears than see it.” (Captain, crew 4).

- “I think I would chose with party line up to a certain point, but when you get the line up for your approach you go without party line then. You don’t really want to have to pay attention to other traffic anymore, I think that would be the best option.” (F/O, crew 3).

The Captain of crew 2 thought that it might be an option to have only the controller speaking and all the aircraft receiving their clearances via data link, so that they do not have to be read back. The First Officer disagreed and believed that this may be even more distracting because you do not get the whole picture, and only hear what is happening from one side. The Captain of crew 3 thought flying without party line would probably be better, but that it may take a while to get used to hearing no R/T other than that for the specific aircraft.

- “You really get a feeling that there’s a communication failure with your aircraft because it’s all of a sudden quiet, but maybe you could have some sort of identification on your radio so you can see it’s still working.” (Captain, crew 3).

It was also noted that attending to data link messages would still cause distraction, and that some pilots may find it worse if they are not confident with the technology.

- “You still have to press buttons to send messages so you’re still distracted. The funny thing is that most of the pilots I know, if they are a little bit older, they get scared, they cannot program computers!”

APPENDIX D: RESULTS OF EEG ANALYSIS

In line with Prinzel (2003) the EEG engagement index was determined for the three measured locations Cz, Pz and Oz. The EEG engagement index is defined as the power of the beta band divided by the power of the alpha and theta band. For presentation purposes, the derived value is multiplied by 100 (as opposed to Prinzel who uses a factor of 20). Since during higher workload periods the alpha and theta band are suppressed whilst the power in the beta band remains at a similar level, the index value should increase. However, the theta band has also been found to be sensitive to memory load. Therefore extensive use of long-term memory may lead to opposite results.

The index was chosen as representative of the EEG results since it is a relative measure. Actual power levels differ considerably between individuals due to the different characteristics of the skull. Measuring relatively within one person removes part of the variability in the results. In addition, the index has shown some interesting results in laboratory experiments, but usage within simulator environments has attracted less attention.

Five 10 s segments (subparts of the 90 s segments as used for the heart rate and blink rate analysis) were identified. EOG artefacts were removed from the EEG signals using a linear regression technique. Since 13 analysis periods were contaminated by strong artefacts caused by factors such as participant movements and measurement deficiencies, these values were replaced by a mean index value taking into account the flight phase and condition, to enable statistical analysis of the data.

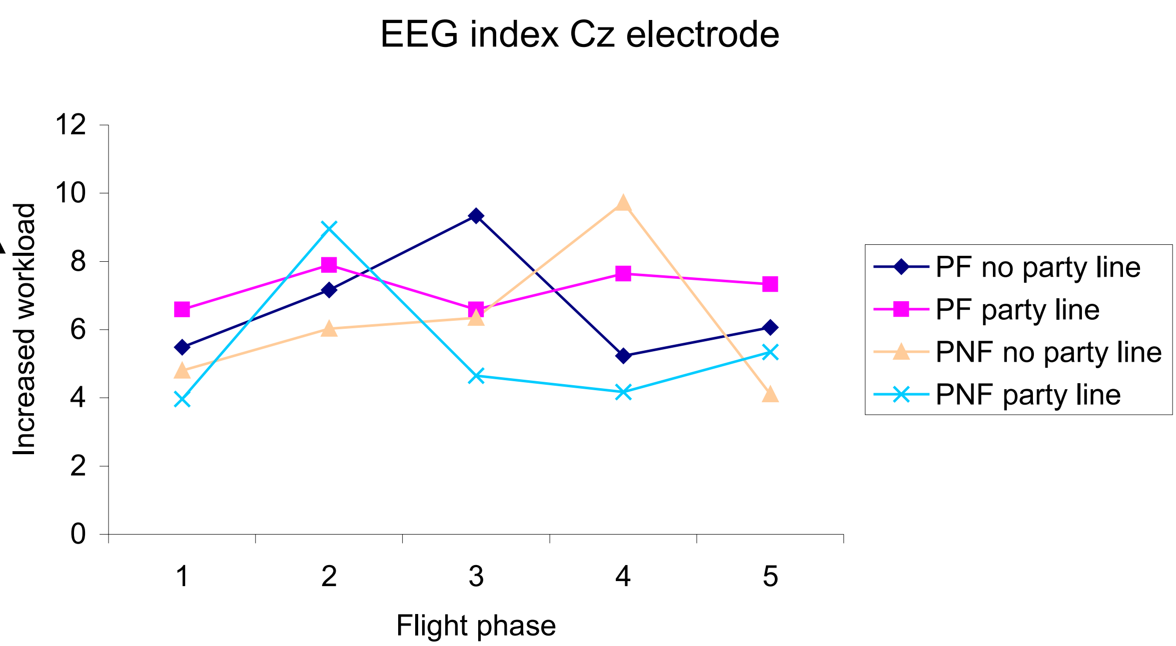


Figure D-1. EEG engagement index for the Cz electrode across phases of flight according to party line condition and flying status. Note: workload increases are indicated by engagement index increases

Figure D-1 shows the EEG index for the Cz electrode channel. A 5 (flight phase) x 4 (flight condition) repeated-measures ANOVA revealed no significant main effect of flight condition, $F(3, 21) = 1.01$, $MSE = 22.78$, $p = .41$, or of flight phase, $F(4, 28) = 1.84$, $MSE = 14.41$, $p = .15$. Also, the interaction between flight phase and condition did not quite reach significance, $F(12, 84) = 1.63$, $MSE = 12.75$, $p = .06$. An analysis conducted separately on data for Captains and First Officers revealed similar null effects.

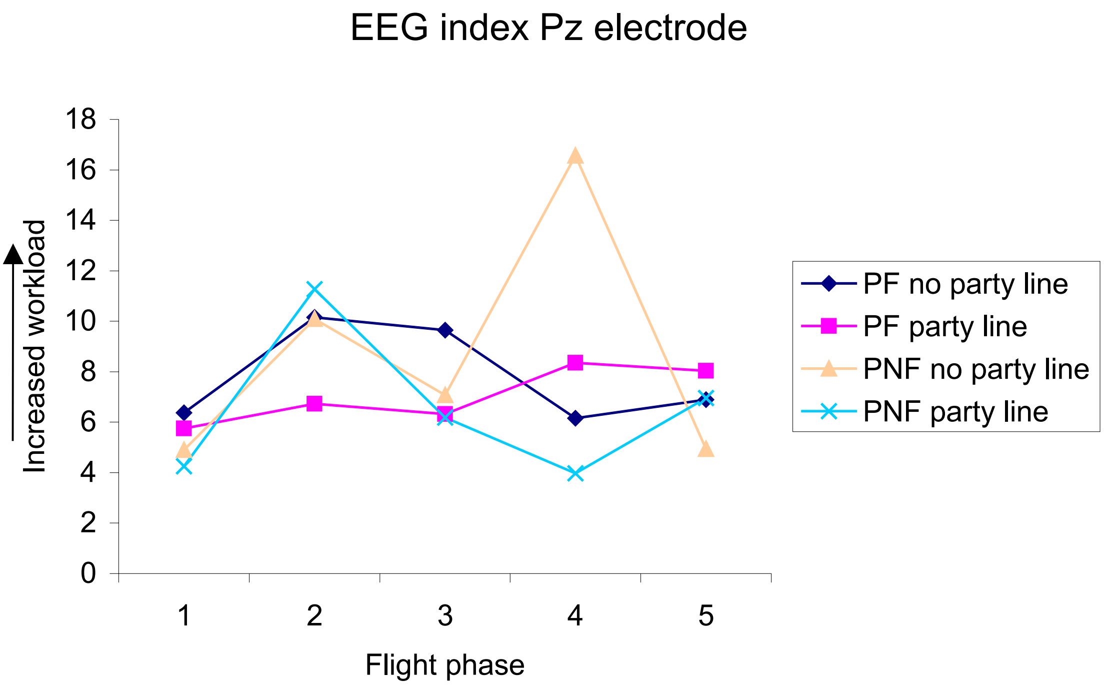


Figure D-2. EEG engagement index value for the Pz electrode across phases of flight, according to party line condition and flying status

Figure D-2 shows EEG index values obtained from the Pz electrode. Again, there were no main effects of flight condition, $F(3, 21) = .63$, $MSE = 58.40$, $p = .61$, or of flight phase, $F(4, 28) = 2.11$, $MSE = 42.55$, $p = .11$. However, there was a significant interaction, $F(12, 84) = 2.08$, $MSE = 33.44$, $p < .03$. Although a number of pairwise comparisons approached significance, the only significant difference at the .05 level was the increase in the EEG index between Flight Phases 2 and 4 in the PF party line condition. No effects reached significance when Captains and First Officers were analysed separately.

Results were perhaps most interesting for the Oz electrode (Figure D-3). Although there was no significant effect of flight condition, $F(3, 21) = 1.19$, $MSE = 56.91$, $p = .34$, there was a main effect of flight phase, $F(4, 28) = 2.68$, $MSE = 58.75$, $p = .05$ such that the EEG index was significantly lower in Phase 1 than in either phases 2 or 4, $p < .02$.

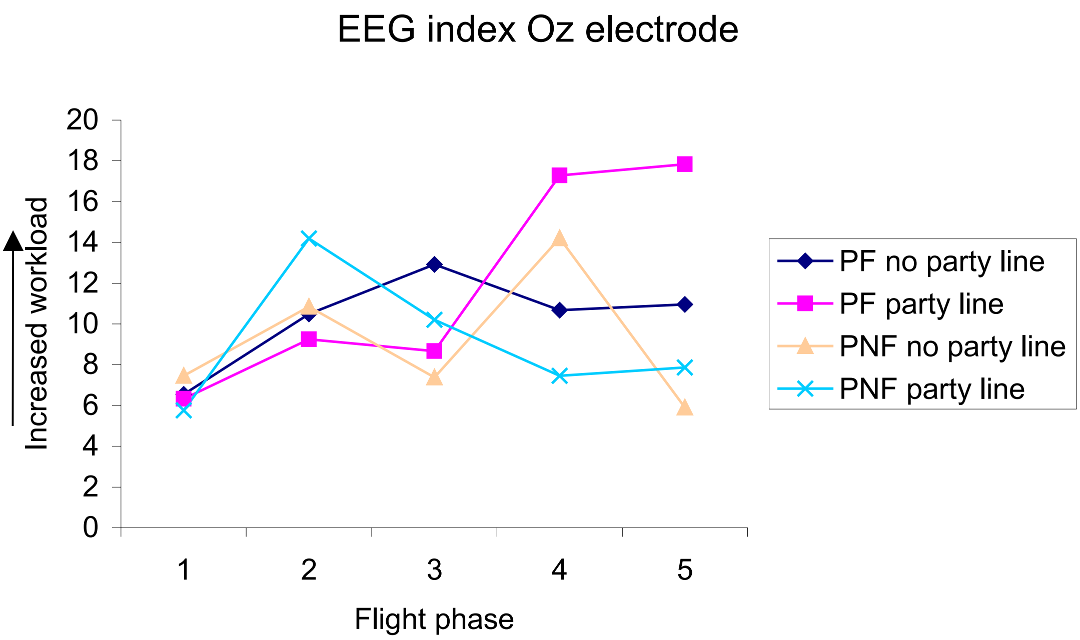


Figure D-3. EEG engagement index values for the Oz electrode across phases of flight, according to party line condition and flying status

There was a highly significant interaction between flight phase and condition, $F(12, 84) = 3.74$, $MSE = 25.50$, $p < .001$. In Phase 3, the EEG index was higher for the PF than the PNF without party line; whereas in Phase 4 EEG levels were higher for the PF than PNF with party line ($p < .04$). Also, the EEG index was significantly higher in the PF party line condition than in either of the two PNF conditions in the final flight phase, $p < .03$. In the PF no party line condition, EEG levels were significantly lower in the first flight phase than in either the second or third ($p < .04$). With party line, EEG activity for the PF was significantly lower in the first and second flight phases than in either Phases 4 ($p < .01$) or 5 ($p < .05$). For the PNF, EEG levels were significantly lower in Phase 1 than 2 ($p < .03$) in the presence of party line, but without party line there were no significant differences between phases of flight.

When analysed separately according to pilot rank, there were no significant effects of either flight condition [Captains: $F(3, 9) = .72$, $MSE = 56.09$, $p = .57$, First Officers: $F(3, 9) = .38$, $MSE = 76.27$, $p = .77$] or of flight phase [Captains: $F(4, 12) = 2.15$, $MSE = 57.31$, $p = .14$, First Officers: $F(4, 12) = .62$, $MSE = 75.54$, $p = .66$]. For First Officers, the interaction between flight phase and condition was highly significant, $F(12, 36) = 5.50$, $MSE = 20.06$, $p < .01$, but this was not the case for Captains, $F(12, 36) = .77$, $MSE = 27.46$, $p = .68$.

General discussion

Only short periods could be selected for the analysis of the EEG data due to the occurrence of artefacts. Therefore in the future more effort should be devoted to the correct measurement of the EEG signals. However it was reassuring that several runs showed relatively clean signals. This is remarkable since the participants were making movements, e.g. to select aircraft controls, etc. Also in electromagnetic terms, an aircraft simulator is a ‘dirty environment’ due to the installed equipment such as computer screens and power supplies.

The fact that the EEG results were inconclusive can be attributed to the low number of independent analysis periods due to the number of artefacts. Comparison of Phase 1 to Phase 2 for the pilot flying shows an expected trend for increased workload after autopilot disconnect.

Alternatively, it was also tested whether it was possible to obtain an Event Related Potential response based on eye-movements. To this end the location of a saccade was determined from the horizontal EOG, and EEG data around the midpoints in time of those saccades were averaged. Similar patterns as Yagi were obtained. It was found that the sensitivity to measurement artefacts was relatively small. However interpretation of those Eye Fixation Related Potentials (EFRPs) with respect to workload is still a topic for further research, and as such the results have been omitted from this report.

APPENDIX E: DETAILED STATISTICAL RESULTS

Probability of interaction of experimental conditions with x and y co-ordinates of the touchdown point ('landing point')

Condition	Probability
Pilot role on x co-ordinate	0.461
Pilot role on y co-ordinate	0.497
Party line on x co-ordinate	**0.022***
Party line on y co-ordinate	0.443
Pilot role x Party line on x co-ordinate	0.073
Pilot role x Party line on y co-ordinate	0.435

Results of analysis of variance for the performance data in accordance with flight condition (with/without party line and flying role PF/PNF) and flight phase

Tests of Between-Subjects Effects

Source	Dependent Variable	df	Mean Square	F	Sig.
Party Line	FD_Pitch	1	.010	.106	.746
	FD_Roll	1	.091	.494	.487
	Pitch_sdev	1	.008	.045	.833
	Roll_sdev	1	.936	.733	.398
Pilot role (Captain or FO)	FD_Pitch	1	.014	.151	.699
	FD_Roll	1	.570	3.087	.087
	Pitch_sdev	1	.429	2.513	.122
	Roll_sdev	1	3.560	2.787	.104
Flight phase	FD_Pitch	2	1.252	13.335	**.000****
	FD_Roll	2	3.097	16.775	**.000****
	Pitch_sdev	2	1.877	11.009	**.000****
	Roll_sdev	2	24.461	19.148	**.000****
Party Line * Pilot role	FD_Pitch	1	.041	.431	.516
	FD_Roll	1	.041	.225	.638
	Pitch_sdev	1	.000	.002	.966
	Roll_sdev	1	.339	.265	.610
Party Line * Flight Phase	FD_Pitch	2	.066	.702	.502
	FD_Roll	2	.028	.152	.860
	Pitch_sdev	2	.230	1.348	.273
	Roll_sdev	2	1.085	.849	.436
Pilot role * Flight Phase	FD_Pitch	2	.197	2.095	.138
	FD_Roll	2	.499	2.702	.081
	Pitch_sdev	2	.174	1.020	.371
	Roll_sdev	2	2.764	2.164	.130

Note: FD_Pitch indicates the standard deviation of the Flight Director Pitch bar command. FD_Roll indicates the standard deviation of the Flight Director roll bar command. The Pitch_sdev indicates the standard deviation of the aircraft pitch angle and the Roll_sdev indicates the standard deviation of the aircraft roll angle.

Results of analysis of variance for physiological data in accordance with flight condition (with/without party line and flying role PF/PNF) and flight phase

Physiological data	Effect	Df	MSE	F	p
Heart rate	Flight condition	(3, 21)	134.05	4.82	**<.01****
	Flight phase	(4, 28)	27.88	14.14	**<.01****
	Flight phase x condition	(12, 84)	9.90	7.32	**<.01****
Heart rate variability	Flight condition	(3, 21)	103.88	3.44	**.04***
	Flight phase	(4, 28)	41.29	4.62	**<.01****
	Flight phase x condition	(12, 84)	41.50	1.30	.24
EOG (blink rate)	Flight condition	(3, 21)	37.27	16.17	**<.01****
	Flight phase	(4, 28)	37.79	8.45	**<.01****
	Flight phase x condition	(12, 84)	17.82	3.17	**<.01****
EEG Cz electrode	Flight condition	(3, 21)	22.78	1.01	.41
	Flight phase	(4, 28)	14.41	1.84	.15
	Flight phase x condition	(12, 84)	12.75	1.63	.06
EEG Pz electrode	Flight condition	(3, 21)	58.40	.63	.61
	Flight phase	(4, 28)	42.55	2.11	.11
	Flight phase x condition	(12, 84)	33.44	2.08	**.03***
EEG Oz electrode	Flight condition	(3, 21)	56.91	1.19	.34
	Flight phase	(4, 28)	58.75	2.68	**.05***
	Flight phase x condition	(12, 84)	25.50	3.74	**<.01****
SPO2 (Captains)	Flight condition	(3, 6)	1.54	2.50	.16
	Flight phase	(4, 8)	9.17E-02	.55	.70
	Flight phase x condition	(12, 24)	8.49E-02	2.52	**.03***
SCL corr (F/Os)	Flight condition	(3, 9)	4.32E-02	.20	.90
	Flight phase	(4, 12)	3.36E-02	11.20	**<.01****
	Flight phase x condition	(12, 36)	4.80E-03	1.35	.23
SCL SD (F/Os)	Flight condition	(3, 9)	1.02E-02	.25	.86
	Flight phase	(4, 12)	4.10E-03	4.50	.02
	Flight phase x condition	(12, 36)	2.87E-03	3.75	**<.01****

Results of analysis of variance for physiological data in accordance with flight condition (with/without party line and flying role PF/PNF) and flight phase, for both Captains and First Officers

			Captains			First Officers		
	Effect	df	MSE	F	p	MSE	F	p
Heart rate	Flight condition	(3, 9)	141.51	3.72	**.05***	143.95	1.40	.30
	Flight phase	(4, 12)	39.19	4.50	**.02***	17.26	14.12	**<.01****
	Phase x condition	(12, 36)	6.44	4.45	**<.01****	14.23	3.59	**<.01****
HRV	Flight condition	(3, 9)	67.68	1.74	.23	161.87	1.72	.23
	Flight phase	(4, 12)	21.89	6.99	**<.01****	60.14	1.36	.31
	Phase x condition	(12, 36)	50.05	1.14	.36	30.12	1.55	.15
Blink rate	Flight condition	(3, 9)	50.09	9.41	**<.01****	21.49	8.25	**<.01****
	Flight phase	(4, 12)	76.16	2.50	.10	5.71	25.92	**<.01****
	Phase x condition	(12, 36)	29.68	1.50	.17	8.24	2.77	**<.01****

Results of analysis of variance conducted on post-flight questionnaire items

Questionnaire item	Effect	F	MSE	p
RSME	Party line	19.02	55.27	**<.01****
	Flying role	12.65	121.24	**<.01****
	Party line x Flying role	2.34	72.21	.17
Mental & perceptual activity	Party line	14.83	80.56	**<.01****
	Flying role	5.82	166.91	**.05***
	Party line x Flying role	5.00	40.54	.06
Perceived time pressure	Party line	6.81	76.96	**.04***
	Flying role	7.53	123.49	**.03***
	Party line x Flying role	4.83	140.63	.06
Success in accomplishing tasks	Party line	2.44	53.37	.16
	Flying role	2.19	415.59	.18
	Party line x Flying role	.47	169.66	.52
Mental workload	Party line	62.94	13.45	**<.01****
	Flying role	19.83	128.43	**<.01****
	Party line x Flying role	.56	91.10	.49
Frustration	Party line	14.08	104.03	**<.01****
	Flying role	1.03	396.82	.34
	Party line x Flying role	.79	323.86	.40
R/T distraction	Party line	9.66	483.14	.02*
	Flying role	3.81	41.92	.09
	Party line x Flying role	1.18	87.48	.31
Perceived performance impairment	Party line	18.43	213.75	**<.01****
	Flying role	.69	99.85	.43
	Party line x Flying role	.06	167.02	.81
Perceived flight safety	Party line	37.22	11.89	**<.01****
	Flying role	.58	248.61	.47
	Party line x Flying role	1.77	133.10	.22

(*p <.05, **p <.01, df (1,7))

Results of analysis of variance conducted on post-flight questionnaire items for both Captains and First Officers

		Captains			First officers		
Questionnaire item	Effect	F	MSE	P	F	MSE	P
RSME	Party line	19.22	38.57	**.02***	4.44	38.57	.13
	Flying role	3.69	194.30	.15	8.49	125.39	.06
	Party line x Flying role	.08	18.07	.79	4.94	77.97	.11
Mental & perceptual activity	Party line	50.43	15.27	**<.01****	2.70	165.39	.20
	Flying role	3.75	205.44	.15	1.63	206.77	.29
	Party line x Flying role	.34	26.29	.60	8.37	35.06	.06
Perceived time pressure	Party line	16.22	17.56	**.03***	1.47	161.71	.31
	Flying role	.50	125.31	.53	32.75	37.94	**.01****
	Party line x Flying role	.01	60.81	.92	76.26	18.69	**<.01****
Success in accomplishing tasks	Party line	9.04	27.89	.06	.001	55.94	.98
	Flying role	.88	306.02	.42	1.06	647.44	.38
	Party line x Flying role	.01	53.39	.91	.59	308.04	.50
Mental workload	Party line	63.27	6.72	**<.01****	17.07	24.63	**.03***
	Flying role	9.98	188.56	**.05***	10.93	71.71	**.05***
	Party line x Flying role	2.84	39.81	.19	37.22	11.29	**<.01****
Frustration	Party line	24.61	29.35	**.02***	3.48	213.35	.16
	Flying role	.59	468.81	.50	.32	453.54	.61
	Party line x Flying role	.08	134.97	.79	1.42	676.00	.32
R/T distraction	Party line	3.31	927.85	.17	10.24	166.23	**.05***
	Flying role	.20	49.06	.69	8.30	26.23	.06
	Party line x Flying role	.01	115.72	.92	2.75	63.90	.20
Perceived performance impairment	Party line	6.01	395.44	.09	17.67	90.54	**.03***
	Flying role	.29	93.94	.63	5.11	56.54	.11
	Party line x Flying role	1.32	143.19	.33	4.39	75.85	.13
Perceived flight safety	Party line	81.42	5.10	**<.01****	35.53	2.47	**<.01****
	Flying role	.07	21.39	.80	2.70	173.06	.20
	Party line x Flying role	.47	28.10	.54	1.33	247.43	.33

(p <.05, **p <.01, df (1,3))

The significant effects are summarised in the following table for Captains and First Officers, together and separately. It can be seen, for example, that:

- only the First Officers reported that they were distracted by the party line
- only the Captains believed that the party line imposed greater time pressure
- both groups reported increase in mental workload

P values obtained for each questionnaire item, showing the effects of party line, flying role and the interaction between these factors. Probabilities are shown for the overall data, as well as for Captains and First Officers separately. Key: * = p < .05; ** = p < .01

Scale	Effect	p (for overall data)	p (captains only)	p (F/Os only)
RSME	Party line	**.003****	**.022***	.126
	Flying role	**.009****	.151	.062
	Interaction	.170	.792	.113
Mental activity	Party line	**.006****	**.006****	.199
	Flying role	**.047***	.148	.291
	Interaction	.061	.600	.063
Time pressure	Party line	**.035***	**.028***	.310
	Flying role	**.029***	.532	**.011***
	Interaction	.064	.918	**.003****
Task success	Party line	.163	.057	.975
	Flying role	.183	.418	.378
	Interaction	.515	.912	.498
Mental workload	Party line	**.000****	**.004****	**.026***
	Flying role	**.003****	**.050***	**.046***
	Interaction	.488	.191	**.009****
Frustration	Party line	**.007****	**.016***	.159
	Flying role	.343	.498	.612
	Interaction	.404	.790	.319
R/T distraction	Party line	**.017***	.670	**.049***
	Flying role	.092	.686	.064
	Interaction	.313	.923	.196
R/T impairment	Party line	**.004****	.092	**.025***
	Flying role	.433	.626	.109
	Interaction	.813	.334	.127
Perceived safety	Party line	**.001****	**.003****	**.009****
	Flying role	.471	.804	.199
	Interaction	.224	.543	.333

ANOVA results Checklist timing

Effect	df	MSE	F	p
Party line condition	(1, 2)	0.13	0.08	0.80
Checklist	(2, 4)	0.31	1.57	0.32
Party line x checklist	(2, 4)	0.01	11.92	0.02*

APPENDIX F: OBSERVATIONS

Instances of ATC calls missed, queried or read back incorrectly, for Party line (P) and No Party line (NP) conditions. Footnotes provide details of calls

	Crew 1	Crew 2	Crew 3	Crew 4	Total
Calls missed (NP)					0
Calls missed (P)		✓[2]			1
Calls queried (NP)	✓[3]	✓[4]		✓[5]	3
Calls queried (P)	✓[6]	✓[7]		✓[8] ✓[9]	4
Incorrect read back (NP)		✓[10]	✓[11]	✓[12]	3
Incorrect read back (P)	✓[13]	✓[14] ✓[15] ✓[16]		✓[17]	5

Comment: the party line information leads to more RT errors than the situation without a party line. The total of RT calls missed or incorrectly read back is 3 for NP and 6 for P. The numbers for calls queried are almost equal.

Instances of checklist items missed or repeated for Party line (P) and No Party line (NP) conditions

	Crew 1	Crew 2	Crew 3	Crew 4	Total
Items missed (NP)	✓[18]				1
Items missed (P)		✓[19]			1
Items repeated (NP)	✓[20]				1
Items repeated (P)	✓[21]				1
Other (NP)				✓[22]	1
Other (P)	✓[23] ✓[24]				2

[2] CS2CP1: “Contact tower 118.1” not read back, ATC repeats the communication 16 seconds later.
[3] CS1FN3: “FL 70” (R/T unclear). Query with each other, then ask ATC to repeat.
[4] CS2FN4: “L HDG 010, speed 180”, query whether speed 160 or 180.
[5] CS4CN6: ask ATC to confirm direct to SPL.
[6] CS1FP6: “R HDG 150 cleared NDB approach”, ask ATC to repeat.
[7] CS2CP5: “R HDG 105, descend 2000ft, QNH 1013, speed 180”. Query with each other HDG 115 or 105 then ask ATC to clarify.
[8] CS4FP1: “R HDG 100, descend 2000ft, QNH 1013, speed 180”. Ask ATC to repeat HDG.
[9] CS4CP3: “Maintain upon reaching.” Asks ATC to repeat.
[10] CS2CN3: “HDG 100, descend 2000ft, QNH 1013, speed 180.” Did not read back heading.
[11] CS3CN2: “After SUGOL, proceed to SPL, speed 250”. Incorrect read back, ATC corrects.
[12] CS4CN4: KLM 204: “Maintaining 2000 ft” ATC: “Confirm 3000 ft” KLM 204: “Maintaining 3000 ft.”
[13] CS1CP2: Read back for wrong aircraft “KLM 1024, after SUGOL proceed to SPL”
[14] CS2CP1: “Contact tower 118.1” read back as 181.1, ATC corrects.
[15] CS2FP2: Read back for wrong aircraft: “KLM 1024 clear to land.”
[16] CS2CP5: “R HDG 155 cleared NDB approach.” PF tells PNF 155 but PNF reads back 115. PF corrects him, and then he reads back again.
[17] CS4CP3: “R HDG 100, descend 2000 ft, QNH 1013.” Did not read back heading, so ATC asks to confirm.
[18] CS1FN3: Landing checklist interrupted: Although resumed, the last item (flaps) was omitted.
[19] CS2CP5: Landing checklist interrupted after speedbrake and resumed at missed approach altitude thus omitting 2 items: landing gear and altimeters (although landing gear down had already been noted prior to the checklist, it was not read and confirmed. Altimeters were not mentioned).
[20] CS1FN3: Landing checklist interrupted after altimeters, altimeters repeated at resumption.
[21] CS1FP4: Landing checklist interrupted after landing gear, landing gear repeated at resumption.
[22] CS4FN2: Approach checklist interrupted before altimeters item confirmed, resumed at prepare for landing and then went back to altimeters.
[23] CS1CP2: Descent checklist omitted.
[24] CS1FP6: Descent checklist omitted.

Instances of items omitted, repeated or resumed incorrectly when checklist is interrupted with an ATC call to the own aircraft

	Crew 1	Crew 2	Crew 3	Crew 4	Total
Resumed at next item	✓[25]✓[26]		✓[27]		3
Item omitted	✓[28]	✓[29]		✓[30]	3
Item repeated	✓[31]✓[32]				2

[25]CS1CN1: Descent checklist interrupted during discussion of approach preparation. Discussion resumed (although completion not confirmed).

[26] CS1FP4: Descent checklist interrupted during discussion of approach preparation. Discussion resumed (although completion not confirmed).

[27] CS3CP4: Landing checklist interrupted after landing gear, resumed at altimeters.

[28] CS1FN3: Landing checklist interrupted: Although resumed, the last item (flaps) was omitted.

[29] CS2CP5: Landing checklist interrupted after speed brake and resumed at missed approach altitude thus omitting 2 items: landing gear and altimeters (although landing gear down had already been noted prior to the checklist, it was not read and confirmed. Altimeters were not mentioned).

[30] CS4FN2: Approach checklist interrupted before altimeters item confirmed, resumed at prepare for landing and then went back to altimeters.

[31] CS1FN3: Landing checklist interrupted after altimeters, altimeters repeated at resumption.

[32] CS1FP4: Landing checklist interrupted after landing gear, landing gear repeated at resumption.