EUROPEAN ORGANISATION
FOR THE SAFETY OF AIR NAVIGATION

EUROCONTROL

This Document is issued as EATM Reference Material. The contents are not mandatory. They provide information and explanations or may indicate best practice.

A Measure to Assess the Impact of Automation on Teamwork

Edition Number	:	1.0
Edition Date	:	29.10.2004
Status	:	Released Issue
Intended for	:	EATM Stakeholders

EUROPEAN AIR TRAFFIC MANAGEMENT

DOCUMENT CHARACTERISTICS

TITLE

A Measure to Assess the Impact of Automation on Teamwork

	EATM Infocentre Reference:	040819-01
Document Identifier	**Edition Number:**	1.0
HRS/HSP-005-REP-07	**Edition Date:**	29.10.2004

Abstract

This report proposes a validated approach on how to assess the impact of automation on teamwork. On the basis of a literature review the latest research results on teamwork and its measurement are described. Using these results, a model for teamwork measurement named '**S**kills, **K**nowledge and **A**ttitudes for **TE**amwork (**SKATE**)' is presented. This model is the basis for a new set of instruments – a questionnaire, an observation form and a self-rating form for the team members. These instruments were validated in a simulation study at the Tower Research Simulator (TRS) of the Dutch National Aerospace Laboratory (NLR). The experimental design and recommendations for the measurements during the study are described, and the results and conclusions of the study are reported.

Work was developed within the 'Solutions for Human-Automation Partnerships in European ATM (SHAPE)' Project conducted, under the supervision of the EATCHIP/EATM(P)* Human Resources Team, by the Human Factors and Manpower Unit (DIS/HUM) of EUROCONTROL, today known as the Human Factors Management Business Division (DAS/HUM).

---

* In February 1999 the 'European Air Traffic Control Harmonisation and Integration Programme (EATCHIP)' was renamed 'European Air Traffic Management Programme (EATMP)'. Since May 2003 it is known simply as 'European Air Traffic Management (EATM)'.

Keywords

Teamwork	Questionnaire
Validation	Skills, Knowledge and Attitudes for TEamwork (SKATE)
Eye tracking	

Contact Persons	Tel	Division
Oliver STRAETER	+32-2-729 5054	Safety and Security Management Business Division (DAS/SSM)
Manfred BARBARINO	+32-2-729 3951	Human Factors Management Business Division (DAS/HUM)

Authors

Jan Roessingh and Rolf Zon

STATUS, AUDIENCE AND ACCESSIBILITY

Status		Intended for		Accessible via	
Working Draft	☐	General Public	☐	Intranet	☐
Draft	☐	EATM Stakeholders	☑	Extranet	☐
Proposed Issue	☐	Restricted Audience	☐	Internet (www.eurocontrol.int)	☑
Released Issue	☑	*Printed & electronic copies of the document can be obtained from the EATM Infocentre (see page iii)*			

ELECTRONIC SOURCE

Path:	G:\Deliverables\HUM Deliverable pdf Library\	
Host System	**Software**	**Size**
Windows NT	Microsoft Word 8.0b	

EATM Infocentre
EUROCONTROL Headquarters
96 Rue de la Fusée
B-1130 BRUSSELS

Tel: +32 (0)2 729 51 51
Fax: +32 (0)2 729 99 84
E-mail: eatm.infocentre@eurocontrol.int

Open on 08:00 - 15:00 UTC from Monday to Thursday, incl.

DOCUMENT APPROVAL

The following table identifies all management authorities who have successively approved the present issue of this document.

AUTHORITY	NAME AND SIGNATURE	DATE
SHAPE Project Leader	O. STRAETER	01.10.'04
Chairman HRT Human Factors Focus Group (HFFG)	V.S.M. WOLDRING	06.10.2004
Manager EATM Human Resources Programme (HRS-PM)	M. BARBARINO	11.10.04
Chairman EATM Human Resources Team / Programme Steering Group (HRT/PSG)	A. SKONIEZKI	12.10.04
Director ATM Programmes (DAP)	G. PAULSON	13.10.04

DOCUMENT CHANGE RECORD

The following table records the complete history of the successive editions of the present document.

EDITION NUMBER	EDITION DATE	INFOCENTRE REFERENCE	REASON FOR CHANGE	PAGES AFFECTED
0.1	01.09.2003		Working Draft	All
0.2	03.10.2003		Draft for HFFG	All
0.3	06.01.2004		Proposed Issue for HRS-PSG Meeting on 28-29 January 2004 (document configuration and editorial changes)	All
1.0	29.10.2004	040819-01	Released Issue (agreed by HRT-PSG Members by correspondence on 15.03.2004) (final document configuration and editorial adjustments)	All

CONTENTS

EXECUTIVE SUMMARY

This report describes an approach to measure team performance in ATM systems. On the basis of a literature review the latest research results on teamwork and its measurement are described. Using these results a model for teamwork measurement (**SKATE** = **S**kills, **K**nowledge and **A**ttitudes for **TE**amwork) was developed. The SKATE Model is the basis for a revised set of instruments – a questionnaire, an observation form and a self-rating form for the team members. These instruments were validated in a simulation study at the Tower Research Simulator (TRS) of Nationaal Lucht en Ruimtevaart Laboratorium (NLR, the Dutch National Aerospace Laboratory). The experimental design and recommendations for the measurements during that study are described as are the results and conclusions of this study.

Work was developed within the ‘Solutions for Human-Automation Partnerships in European ATM (SHAPE)’ Project conducted, under the supervision of the EATCHIP/EATM(P)[1] Human Resources Team, by the Human Factors and Manpower Unit (DIS/HUM) of EUROCONTROL, today known as the Human Factors Management Business Division (DAS/HUM).

In Section 1 a logical structure is introduced shifting from more general teamwork/ automation issues via team tasks to required teamwork skills, knowledge and attitudes, and finally to the question how these skills, knowledge and attitudes can be measured during simulation studies.

Section 2 considers the positive and negative effects of several factors (including automation) on teamwork, with a focus on teamwork in domains other than ATM; teamwork in medicine, offshore oil industry and nuclear power industry are considered.

Section 3 provides an overview of team tasks in ATM. This overview is primarily based on the Integrated Task Analysis (ITA) for en-route control (see EATMP publication ‘Integrated Task and Job Analysis of Air Traffic Controllers – Phase 2: Task Analysis of En-route Controllers’ - EATMP [1999a]).

In Section 4 consideration is given to teamwork Skills, Knowledge and Attitudes (SKAs) that allow for the functioning of the team, i.e. successful completion of team tasks. The results of the analysis of literature findings are encapsulated in the SKATE Model.

Section 5 considers different types of measures for teamwork such as team process measures, team outcome measures, measures which can be obtained from observers, from participating controllers, and from the simulation. This section also proposes revisions to the instruments and recommends methods for analysis of the data collected with the measure.

[1] In February 1999 the ‘European Air Traffic Control Harmonisation and Integration Programme (EATCHIP)‘ was renamed ‘European Air Traffic Management Programme (EATMP)‘. Since May 2003 it is known simply as ‘European Air Traffic Management (EATM)‘.

Section 6 concludes with a plan for the experimental validation of the revised teamwork measure.

Section 7 describes the validation results.

Section 8 discusses the overall approach and concludes further steps.

Details for publications referred to in this report are then provided, as are definitions of abbreviations & acronyms used and acknowledgements to the contributors to this document.

The appendices contain detailed information on the development of the instrument, the material used and the results: Appendix A describes how teamwork is tackled in other domains, while Appendices B, C and D respectively provide the questionnaires applied, the observation form, and the data and results.

1. INTRODUCTION

Generally spoken, team tasks are at considerable risk when automated technologies are introduced. Automation effects operating at the individual level may have a gamut of effects when distributed across teams. As automation entirely or partially replaces team functions, team structure and changes the composition of the team, team roles are unavoidably redefined and communication patterns are altered (Bowers *et al.*, 1993; Wiener, 1993; Jentsch *et al.*, 1995; Bowers *et al.*, 1996; Mosier & Skitka, 1996). While in the past it was assumed that workload would decrease with the introduction of automation, this advantage has been only partly realised. Automation substitutes human activities by 'machine activities' in combination with new human activities, while not leading to lowered workload levels. Additionally, Situation Awareness (SA) may decline as a result of (1) monitoring demands and subsequent vigilance decrements, (2) complacency due to over-reliance on automation, (3) system complexity, (4) poor interface design, (5) inadequate training or (6) lack of trust in automation (Endsley, 1997; Paris *et al.*, 2000).

The 'Solutions for Human-Automation Partnerships in European ATM (SHAPE)' Project initiated in 2000 within the EATMP Human Resources Programme (Human Factors Sub-Programme), which was conducted by the EUROCONTROL's Human Factors and Manpower Unit (DIS/HUM), today known as 'Human Factors Management Business Division (DAS/HUM)' (see EATMP, 2000), addresses the challenges on human factors as a consequence of the introduction of automation in ATM. These challenges concern:

- the level of trust that controllers have in automated tools (see EATM, 2003a, b, c);
- the effect on the controllers' situation awareness when using automated tools (see EATM; 2003d);
- the changes in skills needed to perform the controllers' job (see EATM, 2004a);
- the recovery from system failures when these occur in automated systems (see EATM, 2004b);
- the changes in (mental) workload that result from working with automation (see EATM, 2004c);
- the level of support needed when older controllers make the transition to a system with a higher level of automation than the one previously used (see EATM, 2003e, 2004d);

- the changes in teamwork when a team of controllers make the transition to such a system (the work on the SHAPE Teamwork Measure presented in this document is concerned with this last point).

As the issue of team-working tackled in this report is complex, it was felt necessary to provide definitions for the main topics developed in it.

Team: A distinguishable set of two or more people who interact dynamically, interdependently, and adaptively toward a common and valued goal/objective/mission, who have each been assigned specific roles or functions to perform, and who have a limited life span membership (Salas, Dickinson, Converse & Tannenbaum, 1992).

In Air Traffic Control (ATC) this goal is a safe and efficient control of traffic, in accordance with procedures and agreements. In principle, one could think of an ATC team, as the team consisting of the controllers of a specific sector and, in addition, the controllers of adjacent sectors, the cockpit crew of aircraft under control and possibly other personnel (such as ATC-system maintenance technicians) present in the control room. However, to keep matters simple and because this project concerns teamwork measures that are applicable to a wide range of controller-in-the-loop simulation studies, the team is defined as consisting of the controllers which are together responsible for one specific sector or aerodrome area, only. Although the crew in the cockpit is not considered part of the team in this report, this does not mean that we assume that teamwork is not affected by the interaction between the automation and the crew in the cockpit. It is just that the focus of the to-be-developed teamwork concept is on the team on the ground.

According to the adopted definition, a team in En-route / Area (ACC) control thus consists of the Tactical Controller (TC) (Radar controller) and the Planning Controller (PC). However, In Approach control and in Tower control, controller functions can be flexibly merged or split, depending on situational factors such as traffic load. Also, the staffing of sectors/areas is different in different countries and airports. Therefore, we will not consider the composition of the team in Approach and Tower control in detail. A team in Approach control consists, for example, of Feeder/Departure controller, Arrival controller and Approach planner (the different functions depending on the specific Terminal Control Area [TMA]). A team in Tower control consists, for example, of a Tower Controller, a Start-up Controller and Ground Controller, again very much depending on the specific airport, time of day and other conditions.

Teamwork: We define teamwork as the seamless integration of specific skills, knowledge and attitudes that allow team members to adapt and optimise their performance.

Skills: We define a skill (or ability) as a goal-directed and well-organised behaviour that is acquired through practice and performed with economy of effort. An example of a teamwork skill is a controller’s ability to predict the behaviour of other team members in a developing traffic situation. Such a skill enables the team member to optimally support the team.

In this project we consider a teamwork skill as an individual skill, that is not the skill of the team as a whole. However, when assessing teamwork skills, it may be straightforward to analyse the behaviour of the team as a whole (for example, when analysing team communications), without being able to assess the skill at the individual level. This is not considered problematic, because the goal of the current project is to refine and validate instruments that measure the impact of automation on teamwork, rather than assessing individual skills.

The basis for the acquisition and fine-tuning of teamwork skills is suitable knowledge and attitudes with respect to teamwork.

Knowledge: Knowledge is difficult to define, but generally the following building blocks are recognised: (1) declarative knowledge (facts and concepts), (2) procedural knowledge: procedures and strategies, and (3) conditional knowledge: principles and conditions.

Examples of teamwork knowledge in these different categories are (1) understanding one's own function in the team, (2) knowledge of communication strategies such as ways to give and receive feedback and constructive criticism, and (3) the principles and conditions for creating and retaining a good teamwork atmosphere.

Teamwork knowledge is typically acquired during education, dedicated Team Resource Management (TRM) courses (to be explained later) and similar specialised initiatives. However, knowledge also builds up through operational experience, which provides insight in the operations, procedures and processes and the knowledge to keep track of the situation and to 'read the game'.

Attitudes: Teamwork attitudes are defined as an internal state that influences a team member's choices or decisions to act in a particular way (Cannon-Bowers *et al.*, 1995). Two examples of teamwork attitudes are (1) belief in the importance of teamwork and (2) belief in continuous learning as one of the main functions of the team.

It is thought that both knowledge and attitudes can be partly acquired and/or enriched through dedicated TRM courses.

Automation: The ultimate central concept in this research is automation, which supposedly has an effect on teamwork. When analysing the impact of automation on teamwork, it is convenient to characterise the type of automation that is introduced into the system. To characterise an automated system in ATC is not straightforward. There are many different automation applications, dedicated to specific airspace, system functions (surveillance, communication, flight information, conflict avoidance, long-range planning, etc.) and controller tasks. We will deal in further detail with the automation of controller tasks in the subsequent sections.

It has been argued that future automation should be driven by the philosophy of human-centred automation. The choice of what to automate should be guided by the need to compensate for human vulnerabilities and exploit human strengths. Fitts (1951) already realised that machines (computers) are better at tasks that require (1) speed, (2) memory, (3) consistency, (4) computation, (5) power output and (6) information capacity, while men are better at tasks that require (7) sensing, (8) perceiving and (9) reasoning.

Sheridan and Verplanck (1978) first proposed ten possible levels of allocation of decision-making tasks between humans and computers. More recently, Parasuraman, Sheridan and Wickens (2000) have considered the application of automation to a four-stage model of independent information processing functions:

1. Information acquisition.
2. Analysis.
3. Decision selection.
4. Action implementation.

On this basis Parasuraman *et al*. proposed a revised set of ten levels of automation:

- the computer decides everything and acts autonomously, ignoring the human;
- the computer informs the human only if it (the computer) decides to;
- the computer informs the human only if asked;
- the computer executes automatically, then necessarily informs the human;
- the computer allows the human a restricted time before automatic execution;
- the computer executes the suggestion if the human approves;
- the computer suggests an alternative;
- the computer narrows the selection down to a few;
- the computer offers a complete set of decision alternatives;
- the computer offers no assistance; the human must make all the decisions and actions.

A notion related to advanced automation is that of autonomy. The term autonomy has been introduced to describe the boundaries of decision authority of advanced automation and intelligent decision systems. Autonomy can be defined simply as the capability to make decisions. Thus, autonomy can be considered in terms of freedom to make decisions, considering:

- constraints on decision-making (limitations, rules and regulations);
- decision-making abilities (authority, responsibility, competency);

- capabilities to make different types of decisions (e.g. from resolutions requiring simple choice to resolutions requiring inductive reasoning).

A more specific consideration of automation in ATC will be provided in further sections.

Starting point of the current study is a pre-study proposing several approaches to teamwork and the instruments to measure these (such as NOTECHS and Models for the Analysis of Team Training [MATT]).

The pre-study also proposes a set of three instruments with which the impact of automation can be measured during controller-in-the-loop simulation studies. These instruments - a questionnaire and two observation forms - can be used for evaluation of team processes and performance in experiments and exercises:

- The **Team Process** questionnaire **(TP)**: A questionnaire for the team members (controllers) on perceived team performance after working with new (automated) equipment,

- The **Team Process Quality** (rating) form **(TPQ)**: A behavioural observation scale which requires observers to rate performance using supporting reference material, which is provided with the form, and

- The **Team Process Frequency (TPF)**: A recording form (observation sheet) on which observers may log certain team behaviour at the time of occurrence.

Previous to the experimental validation in an actual simulation study, EUROCONTROL requires to revise or refine this set of instruments in order to achieve a better match with recent findings reported in the literature and theories developed on teamwork, including a consideration of teamwork in non-ATM related domains.

The current report extends the review of the literature that was started in the pre-study. It is further based on a review of task analyses, including EUROCONTROL's Integrated Task Analysis (ITA). The principles derived from literature and analysis are postulated in the **S**kills, **K**nowledge and **A**ttitudes in **TE**amwork (**SKATE**) Model. On this basis of the SKATE Model, revisions of the instruments are proposed.

The review is based on recent research reports (including EATCHIP/EATMP/EATM reports), books and articles (including those appearing in other domains). Where the authors encountered useful older material, this will be cited in the review, in particularly when contributing to the theoretical underpinnings of the teamwork measure.

Page intentionally left blank

2. TEAMWORK IN OTHER DOMAINS

2.1 Introduction

The goal of this section is to examine teamwork in domains that have similarities with the target domain: Air Traffic Management (ATM). Approaches that have been adopted in other domains to manage the introduction of automated tools will be described.

The main question breaks down into the following questions:

- Which models or theories of teamwork are used in other domains?
- What is the impact of automation on teamwork in other domains and how is this assessed?

Most approaches to teamwork in other domains are not directly applicable to ATM, because task characteristics (time scale of problems, task hierarchy, procedures), team composition (team size, spatial distribution of the team), working environment and safety implications are vastly different in every domain. However, the review of teamwork in other domains may help to generate ideas for the validation of teamwork measures in the SHAPE Project, by learning from the practices that have been applied in other domains. An example is the use of Eye-Point-Of-Gaze (EPOG) measurements in addition to measures of communication, which has been investigated for the nuclear industry (Fukuda *et al.*, 2003).

2.2 Nuclear Power Industry

2.2.1 Research material

Teamwork in the nuclear power industry was investigated by reviewing literature (Crichton & Flin, 1999; Fukuda *et al.*, 2003) and an interview (which was held in the context of the 'Enhanced Safety through Situation Awareness Integration in training [ESSAI]' Project [2000, 2003]) with a scientific expert from the UK on emergency management in nuclear power industry. A narrative description of teamwork in the nuclear power industry is included in Appendix A The information sources only describe abnormal (emergency) situations in the nuclear power industry; thus only teamwork under these circumstances will be described.

2.2.2 Summary of findings in the nuclear power industry

Team structure: In case of emergencies, several teams are embedded in the 'response organisation'. These teams are located on site (e.g. the control room) and off site (at the central emergency support centre). Team size varies from two to eighteen persons. Some of these teams are *ad hoc* (team members only work together in case of an emergency).

Team tasks: The task of the 'response organisation' and in particular for the team in the control room is 'to cope with a non-normal event', in particular by controlling the outcome of the event by intervening as early as possible. For example, in an experiment (Fukuda *et al.*, 2003), the level of a reactor pressure vessel starts falling. After the activation of reactor quick shut-down by the designated operator, s/he should check the important variables with the checklist, which in case of problems requires further action and monitoring. Verbal exchange of information with a second operator is of vital importance to successfully cope with this (simulated) non-normal event.

Models and theories of teamwork: Fukuda *et al.* (2003) developed a (communication) model in which certain types of communications (exchange of compatible and incompatible information with the subtask on hand) are coupled to observable actions on the one hand and visual search patterns on the other hand.

Crichton and Flin (1999) focus on different styles of decision-making, depending on the level of command, time pressure and available information. Their approach to individual decision-making (but within the team context) is couched within Rasmussens' Model of Skill-, Rule-, and Knowledge-based behaviour.

The role of automation/technology: Systems in Nuclear Power Plants (NPPs) have such a high level of complexity that it is acknowledged that not all possible non-normal states can be predicted and anticipated beforehand. Therefore, different coping strategies must be used: (1) if the fault underlying the non-normal state is recognised, then so-called Station Operating Instructions (SOIs, rigid procedures) are used to cope with the non-normal state; (2) if the non-normal state is characterised by a number of symptoms, but no specific diagnosis can be made based on technical knowledge, then so-called Symptom-based Emergency Response Guidelines offer general guidelines and advice to cope with the non-normal event; (3) if the non-normal state is potentially threatening to the larger civil public, then Severe Accident Guidelines (SAGs) come into place. In this case, it is not advantageous to further analyse the symptoms of non-normal state, but action should be taken to shut down (parts of) the plant.

Teamwork measurement techniques: Fukuda *et al.* (2003) report use of verbal protocols and Eye-Point-Of-Gaze (EPOG) measures. EPOG is measured in order to analyse the quality of the team interaction, at the assumption that the eye movement pattern is influenced by the communication behaviour. If the verbal information given to team member whose EPOG is measured matches with the current subtask, then the eye movement pattern does not change. However, if the provided information does not match with the current subtask, then the operator will search for more relevant information and as a result, the eye movement pattern will change. Further, an interviewing technique is used in which verbal retrospection is enhanced by confronting team members with the eye-tracking video (a video that displays the area of interest with fixation points).

Training solutions: In the UK (Crichton & Flin, 1999) training in Team Resource Management (TRM) and Tactical Decision Games (TDGs) are used. TRM is a concept similar to Crew Resource Management (CRM) in aviation. Its most general definition is: the effective use of all resources available to the team in order to achieve safe and efficient operations. In the current context, this operation is the management of an emergency in an NPP.

TDGs offer a variety of scenarios, differing in complexity, timescale and technicalities. The goal of this relatively new training method is to exercise team decision-making and illustrate key operating principles. TDG should enable team members a shared understanding and recognition of possible problems for emergency management.

2.3 The Medical Domain (Anaesthesia)

2.3.1 Research material

This section is mainly based on interviews (in the context of the ESSAI Project [2000, 2003]) held with a senior anaesthetist at a regional hospital, with an anaesthetists' CRM researcher at the University of Aberdeen in Scotland. The material was delivered by Gaba, Fish and Howard (1994), Flin (1996), and a literature review was provided by Fletcher *et al.* (2002). A narrative description of teamwork in anaesthesia is included in Appendix A

2.3.2 Summary of findings in anaesthesia

Team structure: A core anaesthetic team consists of one anaesthetist and an anaesthetic nurse, but can be extended with extra staff of either type, as special circumstances require. The team works inside the operating theatre at the hospital.

Team tasks: The tasks in the operating theatre are essentially threefold. Firstly to ensure that that the patient is unaware and cannot recall any of the surgical stimuli; secondly, to provide conditions that will allow the surgeon to carry out the operation; thirdly, to ensure that the patient 'experiences as little morbidity as possible'. At the hospital that was visited, ca. 20% of all operations are emergency operations, which come in through the Emergency Room (ER). These operations need to be carried out within 24 hours. Within this group, 5% of the operations need to be carried out within half an hour (the team has such an operation once in every two days, on average). These operations are all considered standard tasks, although extra alertness for sudden mishaps, such as sudden blood-loss, is necessary in the group that comes in via the ER. In these cases, no pre-operative assessment of the patient is possible.

Models and theories of teamwork: The predominant model for teamwork in anaesthesia in the US and UK (which is the origin of the material studied for the current domain) is Helmreich's error management model. For decision-making, Gaba *et al.* (1994) hypothesise a classification of which the first three levels correspond to Rasmussen's hierarchy of skill-based, rule-based and knowledge-based behaviour: (1) sensorimotor, (2) procedural, (3) abstract reasoning, (4) supervisory control (allocation of attention, prioritisation, scheduling and coordination of action), and (5) resource management (command and control of available resources, basically CRM with a focus on communication, delegation and coordination of activities).

The role of automation/technology: Like in ATC continuously more confidence is put in technology in anaesthetics. Elder anaesthetists used to do small repair jobs when their equipment had certain failures. Nowadays, only the 'self-test function' of digital equipment can be operated in case of (suspected) malfunction. Basic standard equipment consist of anaesthetic apparatus, consisting of (electronic) breathing apparatus and monitoring apparatus, checking heart, blood-pressure, oxygen saturation, CO_2 and temperature. Other standard apparatus include an infusion apparatus, pumps for medication and heat mattresses. Less standard equipment is a bi-spectral analysis apparatus, to measure the depth of the anaesthesia, via Electro-EncephaloGram (EEG).

Teamwork measurement techniques: Flin *et al.* (2003) measured, among other things, attitudes towards teamwork. These researchers undertook a questionnaire survey with 222 anaesthetists from eleven Scottish hospitals to measure their attitudes towards human and organisational factors that can have an impact on effective team performance. The researchers used the Operating Room Management Attitude Questionnaire (ORMAQ), which is based on the Flight Management Attitude Questionnaire (FMAQ, developed for NASA by Helmreich and his colleagues, see e.g. Sherman & Helmreich, 1995) to measure the attitudes of pilots towards automation.

The questionnaire measures attitudes towards leadership, communication, teamwork, stress, fatigue, work values, human error and organisational climate. The results reveal (among other things) that 'some anaesthetists do not fully appreciate the debilitating effects of stress and fatigue on performance'.

Training solutions: From the literature review by Fletcher *et al.* (2002) it was learned that in the US Anaesthesia Crisis Resource Management (ACRM), training was devised to help address the problem of limited exposure to emergency situations. Following on from CRM training in aviation a course was developed for anaesthetists to help them develop 'pre-compiled responses to critical incidents' and 'to instruct participants in the coordinated integration of all available resources to maximise safe patient outcomes'. More information on this ACRM course is included in the appendices. It should be noted that a current obstacle for assessing the effectiveness of ACRM training is the absence of a validated performance measure for resource management behaviours.

2.4 Offshore Oil Industry

2.4.1 Research material

This section is based on material received from an expert on emergency management in Europe. Additionally an interview was carried out with this expert, who has done extensive research on crisis management and incident command in the offshore oil industry. The researched material includes books by Flin (1996) and Flin and Slaven (1994), and a paper by Flin, Slaven and Stewart (1996). A narrative description of the domain is included in the appendices. As most of the material deals with teamwork in emergency situations (particularly the disaster with the Piper Alpha), only teamwork in these specific situations will be considered.

2.4.2 Summary of findings in the offshore oil industry

Team structure: Several hundred people are on board an oil and gas production platform and fall under the command of the Offshore Oil Installation Manager (OIM). For example, on board the Piper Alpha at the time of the disaster (6^{th} of July, 1988) 226 people were on board. However, the platform emergency response team is much smaller. The team consist of the OIM of the platform and more senior managers who are based onshore. In addition, OIMs of neighbouring platforms take part in the emergency response team. In the event of a serious offshore incident, an onshore emergency response team is gathered in the company's shore-based office, some of which have dedicated emergency control centres. This can create a distributed decision-making situation, with onshore managers and specialists becoming involved.

Team tasks: In case of an emergency on the platform, the tasks of the emergency response team are diverse:

- giving instructions and guidance to personnel on board the platform;
- decision-making with respect to shutting-off hydrocarbon pipelines, for example those supplying hydrocarbons from neighbouring platforms;
- shutting down oil/gas production;
- mustering personnel;
- deploying fire-fighting or rescue teams;
- liaising with adjacent installations, onshore management, the coastguard, shipping and aviation;
- evacuation of personnel;
- total abandonment of the installation.

Models and theories of teamwork: We identified no specific teamwork models or theories in the domain. However, Flin (1996) puts much emphasis is on distributed decision-making (i.e. team coordination). Flin notes that *the interface of the decision approaches of commanders at different levels – operational, tactical and strategic – raises a number of interesting research issues. Moreover, when this is compounded by distributed decision-making (where those involved in the decision-making are located apart, often by hundreds of miles), then the possibility of decision clashes begins to increase.* This may have relevance for distributed decision-making in ATM, e.g. in emergency situations at the aerodrome, possibly requiring response from emergency troops, (potential) crises requiring unusual inter-sector negotiations, etc.

The role of automation/technology: No specific information. It was mentioned that in an emergency situation the onshore/offshore interface caused problems: the onshore managers failed to realise the severity of the situation, questioned the OIM's decisions, needlessly rearranged evacuation arrangements resulting in serious confusion, asked for information rather than giving it, or did not know enough about the installation.

Teamwork measurement techniques: No specific information. Flin, Slaven and Stewart (1996) carried out interviews with OIMs to investigate aspects of team coordination.

Training solutions: In a CRM course for offshore control room operators' emergency response training, a communication module covered the basic communication process, barriers to effective communication and awareness of strengths and weaknesses in personal communication skills. Exercises highlighted the importance of feedback and listening skills, the role of non-verbal communication and effective communication techniques. An actual offshore incident involving a communication problem was also presented and discussed.

A number of organisations now offer specialist courses, such as Emergency Management Command and Control (EMC[2]) or Offshore Command Training Organization (OCTO). These courses draw on the expertise of former naval commanders (including, in the latter case, two admirals from the Falklands conflict) to provide intensive coaching and development of managers, using both outdoor training and emergency command centre simulator exercises. The oil companies also conduct regular training on their installations, from weekly lifeboat drills to full-scale exercises.

2.5 Tentative Recommendations Following from the Investigation of Teamwork in Other Domains

The initial question regarding teamwork in other domains was broken down into the following two sub-questions:

- Which models or theories of teamwork are used in other domains?

- What is the impact of automation on teamwork in other domains and how is this assessed?

Obviously, we found large differences between the three investigated domains and ATM. However, the investigation yielded two interesting approaches that may be of use in the current project.

Attitudes: It was found in the medical domain that attitudes towards teamwork are included in the teamwork concept and that an effort has been made to measure these attitudes – also in relation to automation. The measurement instrument for measuring attitudes is the Operating Room Management Attitude Questionnaire (ORMAQ). Examples exist of similar questionnaires in other domains that are more directed towards teamwork and the use of automation (such as the Flight Management Attitude Questionnaire [FMAQ]).

Communication and eye-movement patterns: In the nuclear domain a technique was developed in which exchange of compatible and incompatible information with the subtask on hand is related to the visual search patterns of the receiver of this information. These visual search patterns can be analysed with Eye-Point-Of-Gaze (EPOG) measures, using a task-oriented area-of-interest technique, and thus provide information concerning the quality of team communication.

Procedures and guidelines: In the nuclear industry a diversification of procedures was found: (1) Standard Operating Instructions (SOIs), (2) Symptom-based Emergency Response Guidelines (SBERGs) and (3) Severe Accident Guidelines (SAGs). SOIs consist of standard operating procedures and checklists, while SBERGs allow some inductive reasoning on behalf of the team, given some well-defined symptoms of non-normal events which can be handled by the team. Finally, for more serious events SAGs come into place.

Dedicated training solutions: The nuclear and oil offshore industries have adopted special courses for team training. Apart from the Crew Resource Management (CRM) courses, Tactical Decision Games (TDGs), courses for Emergency Management Command and Control (EMC[2]) and Offshore Command Training Organization (OCTO) have been developed.

It should be noted that the interviews and review of research articles do not fully reflect trends in teamwork and the impact of automation within the three investigated domains. The investigation was too limited for that purpose.

Page intentionally left blank

3. AUTOMATION AND TEAM TASKS IN ATM

3.1 Introduction

In the introduction of this report (see Section 1) we defined teamwork as the seamless integration of specific skills, knowledge and attitudes that allow team members to adapt and optimise their performance. By optimising performance we mean optimising performance for the tasks assigned to the team on the one hand and its individual members on the other hand. This suggests that we can distinguish between individual tasks and shared (team-) tasks. However, shared tasks can be broken down into parts which can be performed by individual team members. Some of these individual tasks must possibly be performed in coordination with individual tasks of other team members. These tasks will either be performed with overlap in duration (in parallel) or with no overlap in duration (in a serial fashion) of tasks of other team members. Other parts of team tasks performed by an individual will not require coordination with other team members. The ‘start – and stop conditions’ for these tasks are triggered by other events than team coordination (i.e. independently of team coordination).

We define a task as a function of the ATC system that is assigned to one or more air traffic controllers. When the function is assigned to the automation we call this ‘a machine-function’, or simply ‘a function’, rather than a task.

A team’s organisation, as well as its ability to adapt to various system loads, is largely a function of the dependencies between tasks. These task dependencies influence how team members interact. We previously assumed that tasks may be performed independently, sequentially, or in parallel with tasks of other team members. A simple example of a team task with merely serial relationships is a relay race (each team member starts to run after receiving the baton from the previous runner). Serially structuring a team will expose it to the distinct possibility of system overload since its performance is determined by the weakest link in the chain (Meister, 1985). A simple example of a team task with merely parallel relationships is tug-of-war. The integrated task and job analysis of en-route air traffic controllers (see EATMP, 1999a) reveals that few tasks of air traffic controllers are executed independently and that much of the teamwork is serially structured (i.e. between different controllers in the same sector and between controllers in different sectors).

When task/operator dependency is high, the organisation of the teamwork is important, i.e. inadequate performance at one position will have fundamental effect on other controller positions. In this sense, the organisation of the team determines team output.

In the current section we consider the team tasks of air traffic controllers and how these are affected by automation. In the next paragraph we review some of the literature on the classification of automation, i.e. automation levels. In the paragraph thereafter we define the functions of air traffic control systems and briefly consider which functions can be assigned to the team such that automation of these functions would affect teamwork.

In the subsequent paragraph we describe the impact of automation on team tasks. Since it is not within the scope to give complete team task descriptions for all ATC functions and all facilities (ACCs, Oceanic control, etc.) we cover the impact of automation on team tasks by means of illustrative examples of automation, which are:

- datalink,
- Electronic Flight Strips (EFSs),
- Converging Runway Display Aid (CRDA),
- Highly Interactive Problem Solver (HIPS), and
- Collaborative Decision-Making (CDM).

3.1.1 Levels of automation, two models explained

In the Introduction we referred to a classification by Parasuraman *et al.* (2000) in ten levels of automation relating to decision and action selection. At the lowest level of automation (manual operation), a specific function (e.g. the planning of clearances) is continuously performed by the human operator, with no machine control. At the highest level (full automation) all aspects of the function (including its monitoring) are delegated to a machine such that only the result of the function becomes available to the human operator.

In their report on the future of air traffic control a panel of the US National Research Council (NRC, 1998), chaired by C.D. Wickens, recommended that for system functions with relatively little uncertainty and risk a high level of automation is appropriate.

However, when the system function is associated with greater uncertainty and risk, the level of automation should not be more than ‘the computer suggests an alternative’ (i.e. the fourth lowest level of the ten-level scale).

The panel adds to this recommendation:

> *Any consideration for automation at or above this level must be designed to prevent: loss of vigilance, loss of situation awareness, degradation of operational skills, and degradation of teamwork and communication. Such designs should also ensure the capabilities to overcome or counteract complacency, recover from failure, and provide a means of conflict resolution if loss of separation occurs.*
>
> NRC (1998)

Thus the next highest level of automation, which is inappropriate (according to the NRC Panel) for systems with greater uncertainty and risk, is ‘the computer executes the suggestion if the human approves’.

The NRC Panel also proposes additional scales for levels of automation, namely scales for two different dimensions of automation:

1. Information automation (information acquisition, integration and implementation).

2. Action implementation.

The absence or presence of the following automation features determines the automation level on the first dimension (information automation):

(a) Filtering.
(b) Information distribution.
(c) Information transformation.
(d) Confidence estimates.
(e) Integrity checks.
(f) Flexible information based on requests from users.

An ATC system function (for example surveillance) that has all of these features is thought to have the highest level of information automation.

The second dimension (action implementation) has only two levels:

(a) Manual implementation.
(b) Automatic implementation.

The NRC Panel recommends that future automation developments should focus on information automation rather than action implementation.

In the part of the SHAPE Project related to skill set changes, a form has been developed to classify the Level Of Automation (LOA) of a component (application or tool) of an ATC system. In contrast to the more general LOAs adopted by the NRC Panel, these LOAs are more specific for ATC systems.

The system component is evaluated on six dimensions, listed below:

(a) Information extraction.
(b) Information integration.
(c) Information comprehension.
(d) Decision/choice.
(e) Response execution.
(f) Information retention.

Clarification of the six dimensions:

1. The automation features that may facilitate ‘information extraction’ are automatic highlighting, cueing, de-cluttering and filtering.

2. The automation features for ‘information integration’ are automatic arranging and prioritisation.

3. The automation features for ‘information comprehension’ are automatic comparison, diagnosis, prediction and testing.

4. The automation features for ‘decision/choice’ are automatic option generation, option prioritisation, evaluation of options and option choice.

5. The automation features for ‘response execution’ (or ‘action implementation’) are automatic input (e.g. voice recognition), output (e.g. speech synthesis), implementing a response and implementing an emergency response (the system judges, according to known rules, that an action/response is required).

6. The automation features for ‘information retention’ are automatic reminders, history tracking and auto-delete.

When all of these features are present, the LOA of the component is the highest possible. The evaluation form for the LOA is part of the SHAPE Toolkit, and definitions of each of the automation features listed above is included in the SHAPE Toolkit.

3.1.2 Functions of the ATC system assigned to teams

This section is based on analyses of tasks and functions as reported by three different project teams (EUROCONTROL SHAPE and ITA Teams and the US NRC (1998) Panel on human factors in air traffic control automation).

3.1.2.1 Functions distinguished in SHAPE

In its present form the SHAPE Toolkit distinguishes the following eighteen ATM functions:

1. Coordinate airspace reservations / non-standard flights.
2. Coordinate with other agencies.
3. Formulate and issue ATC clearances.
4. Maintain and update flight data information.
5. Manage degraded systems.
6. Monitor the present and forecast weather.
7. Plan and control aircraft within the area of responsibility.
8. Provide a service in emergency and unusual situations.
9. Provide flight information.
10. Provide a service within defined boundaries.
11. Work as a team member.
12. Operate communication equipment.
13. Provide area procedural service.
14. Cooperate with Military (MIL) ATC.
15. Provide area radar service.
16. Provide the pilot with relevant information.
17. Monitor availability and serviceability states.
18. Operate radar equipment.

One can make a distinction between normal (or routine) functions, such as ‘maintain and update flight data information’ and non-normal (or emergency) functions, such as ‘manage degraded systems’. When the team has to engage in non-normal functions, the potential for human error strongly increases, particularly when the non-normal functions are highly automated. For example, in non-normal situations involving functions assigned to both automation and humans, it may be more difficult to call for assistance from a supervisor. If some controller tasks are performed by automation, the ongoing responsibilities (or roles) of the man and machine may be unclear to the supervisor (CAST, 1998b).

3.1.2.2 Functions distinguished by the NRC Panel

The NRC (1998, pp. 69-84) Panel on human factors in air traffic control automation distinguishes the following top-level functions:

- plan/resolve,
- predict longer term,
- compare, predict very short term,
- transmit information,
- remember,
- identify.

NRC also provides a more fine-grained analysis of sub-functions, on the basis of these top-level functions. We omit this fine-grained analysis but provide an overview of team tasks (or tasks that are potentially team tasks) that can be identified from the analysis.

Share plans with team members, which concerns plans such as:

- special-use airspace activities,
- traffic flows,
- responses to weather, emergencies and system failures, and
- search for lost or overdue aircraft.

Share observations and predictions with team members, such as:

- equipment and system problems,
- predicted loss of separation,
- predicted conflict information,
- predicted deviations from trajectories,
- predicted weather, and
- predicted capacity.

Assist in remembering (or making aware of) objects and events such as:

- tracks of aircraft,
- flight plans and updates,
- conflict situations,
- non-controlled objects,
- assigned aircraft,
- sectorisation, and
- weather information.

Assist in identifying (point-out) items and features such as:

- navigation fixes,
- weather features,
- borders of special-use airspace,

- aircraft parameters,
- aircraft type/designation,
- non-controlled objects,
- ground hazards.

3.1.2.3 *Functions distinguished in ITA*

In the Integrated Task Analysis (ITA) (see EATMP, 1999a) cognitive core tasks (i.e. mental processes) of Air Traffic Controllers (ATCOs) have been well defined:

- maintaining SA (keeping ahead of the traffic, pre-planning),
- making decisions for control actions,
- developing and updating a sector traffic plan.

Although the ITA is a cognitive task analysis, not specifically concerned with teamwork, a fine-grained analysis of the core tasks within ITA reveals the following team tasks:

- An important team task is taking over the working position from another controller. Under normal working conditions the mental picture is built up when taking over the position. Maintaining Situational Awareness (SA) is only possible if a mental picture has adequately been built up. When SA is maintained a continuous anticipation of future situations is possible.

- Other team tasks mentioned in the ITA include: to provide assistance in abnormal situations, complementary tasks such as briefing, continuous checking of the whole situation for abnormal and unexpected events, which includes team monitoring.

The ITA also addresses team flexibility, for example to manage workload:

- Managing team workload: In en-route control units where planning or coordinating controllers are present only in high traffic load situations, one of the first steps for the radar controllers is to call for a planner or coordinator. They take over part of the work, e.g. the Planning Controller (PC) takes over the pre-planning of the traffic to avoid conflicts based on proper long-term traffic planning, so that the radar controller can concentrate on the actual situation in the sector. Another step may be to split the sector.

- For managing ‘routine traffic’ in Approach control (insofar one can speak of ‘routine traffic’ in Approach control) there is flexibility in team functions – Some of the functions (at Schiphol Terminal Control Area [TMA], e.g. feeder, departure controller, arrival controller and approach planner) can be flexibly combined or split, depending on traffic load.

- Conflict detection in en-route control depends (among other things) whether there are PCs or not and it therefore depends on the traffic load. If there are planners, they do the pre-screening for conflicts. Then, they take either action by themselves, e.g. they coordinate a solution with the previous sector (telling or showing the radar controllers afterwards), or they make the radar controller aware of the problem. They discuss together to try to find a solution, or the problem is left to the radar controllers and they decide what to do. Smooth and efficient teamwork requires the radar and PCs to have a tacit [implicit] understanding.

- With respect to attitude / working style: En-route controllers prefer to work with colleagues with colleagues who have a similar working style and similar attitudes towards providing service and who prefer similar solutions to give problems. Tacit understanding is linked with a shared mental model and shared anticipations. Teamwork goes more smoothly if controllers are able to anticipate each others' reactions and what the others accept or are capable of.

ITA also addresses team communication in the Tower:

- Owing to time constraints of their task, aerodrome controllers usually make final decisions about requests instantly. Coordinations with pilots and colleagues about alternative solutions in the event that a request cannot be approved, are rather rare.

Generally spoken, teamwork in aerodrome control (and to a lesser extent approach control) is performed in a even more serial fashion than in en-route control, which means that one team member acts on an aircraft after the previous team member. In en-route control, team members (tactical and planner) operate in parallel to some degree and consequently there is more need and more opportunity for cooperation and coordination. This is also possible because of the wider time-window for the resolution of problems.

As a summary of this Section, Table 1 presents ATM functions that can possibly be assigned to a team to controllers (i.e. team tasks). Where possible, a distinction between normal and non-normal team tasks is made. Each of these team tasks can, in turn, be assigned to one or more higher level ATM functions, using for example the following classification:

- surveillance/identification,
- communication,
- (transmit) flight info,
- conflict avoidance,
- conflict resolution,
- long-range planning,
- remember.

Table 1: ATM functions possibly assigned to a team of controllers

Normal (routine) team tasks	**Non-normal (emergency) team tasks**
1. Coordinate airspace reservations/non-standard flights.	18. Coordinate with other agencies in case of non-normal situations.
2. Routine coordinations with other agencies.	19. Manage degraded systems – add/ask for human support to tasks normally accomplished by a single controller.
3. Maintain and update flight data information.	20. Monitor availability and serviceability states.
4. Monitor the present and forecast weather.	21. Share plans with respect to responses to emergencies, unusual weather conditions, system failures and search for lost or overdue aircraft.
5. Plan aircraft within the area of responsibility.	22. Make others aware of unusual tracks of aircraft, unusual weather conditions, unusual conflict situations, and unusual non-controlled objects.
6. Control aircraft within the area of responsibility.	23. Assist in identification of non-controlled objects and ground hazards.
7. Provide flight information.	24. Taking over the working position from another controller under non-normal working conditions.
8. Cooperate with MIL ATC.	
9. Monitor availability and serviceability states.	
10. Share plans with respect to special-use airspace activities, traffic flows and responses to weather.	
11. Make others aware of tracks of aircraft, flight plans and updates, conflict situations, non-controlled objects, assigned aircraft, sectorisation, and weather information.	
12. Assist in identification of navigation fixes, weather features, borders of special-use airspace, aircraft parameters, aircraft type/designation, non-controlled objects and ground hazards.	
13. Taking over the working position from another controller under normal working conditions.	
14. Engage in (de-)briefing.	
15. Monitor fellow team members for performance, SA workload.	
16. Engage in the splitting and merging of functions, sectors or areas to cope with team task load.	
17. Anticipate each other's reactions, capabilities and acceptance-levels.	

3.2 Impact of Automation on Team Tasks

In this section we provide some examples of the impact of automation on team tasks. To this end we distinguish main subsystems of the ATC system, which are surveillance/identification, communication, (transmit) flight information, immediate conflict avoidance/resolution, strategic long-range planning, and remembering of information. We provide examples of automation in each of those subsystems, with the exception of surveillance and remembering. We assume surveillance technologies, such as radar-based systems, to be present in current systems. We refer to NRC (1998) for a consideration of the future developments in surveillance systems. Some of the examples were taken from EUROCONTROL and European Commission research studies, such as:

- ‘Operational Display and Input Development (ODID)’ IV,
- ‘Programme for Harmonised ATM Research in EUROCONTROL (PHARE)’ Demonstrations - PD/1 through PD/3,
- and ‘Consequences of future ATM systems for ATCO Selection and Training (CAST)’,

complemented with information from the US ‘National Research Council (NRC)’ Study.

In addition, the ATC system has support functions such as training and maintenance. Undoubtedly, automation also impacts teamwork in these areas, however these are not within the scope of the current project.

3.2.1 Automation in communication systems

A communication channel (e.g. digital datalink) can by defined by the medium and quality through which team members engage in their interactions. Obviously, the most common communication channel (or modality) is face-to-face verbal interaction. Other modalities relevant to ATM include paper-mediated (flight strips, telex, fax, etc.), audio-mediated (telephone, Radiotelephony [R/T]) and automation-/computer-mediated communication. All team processes may all be differentially impacted by these modalities.

Automation-/computer-mediated communication tends to:

(1) Limit severely the social context cues (visual feedback and status cues) that are available in face-to-face communication.

(2) Constrain the depth of discussion and analysis that is likely to occur.

(3) Increase the time needed for groups to make a decision.

On the other hand, computer-mediated interaction tends to mask status differences, resulting in greater participation by members and possibly better team coordination (Urban *et al.*, 1995). This effect has also been observed in studies into cockpit automation: a highly automated cockpit seems to mask the ‘power-distance’ between captain and first-officer.

Example application: datalink

Datalink is a set of technologies designed to relay communications between ground and air, using digital information rather than conventional radiotelephone communication channels (Kerns, 1994).

Datalink itself is a form of computer-based automation, but within the datalink system various higher levels of automation have been proposed. Various forms of computer-based automation can assist in message composition and message gating (passing the message into the aircrafts’ flight management system onboard the aircraft).

In the USA simulations (FAA, 1996; NRC, 1998) have revealed the positive benefits of datalink for teamwork on the ground. It enables load sharing and the flexibility of distribution of responsibilities when traffic load becomes quite high. Unlike the R/T of the Tactical Controller (TC) in conventional systems, a datalink can allow various operators to assume temporary responsibility for certain aspects of communication. In simulations, this flexibility has been found to provide an unexpected benefit to control efficiency. The NRC (1998) Panel on human factors on control automation notes that the flexibility can have its downside, unless Team Resource Management (TRM) training is carefully implemented, so that shifts in responsibilities are clearly and unambiguously announced.

In the experimental trials with the EUROCONTROL PHARE PD/1 (1997) system it was attempted to try to maximize the intuitiveness of the datalink command interface via a message-out window in which predefined messages (which were themselves the output of an automated planning process) could be easily selected and then uplinked. It was found that the approach reduced (team) workload. Down-linked messages were also presented visually, in the message-in window, which was as close as possible to the radar display (or Plan View Display [PVD]), such that the controller did not have to gaze away from the PVD.

The potential of digital datalink can be considered with respect to the highly accurate 4D prediction and guidance support it can provide as well as providing relief for the regularly high R/T communication load put on controllers (e.g. Porterfield, 1997). However, controller performance in evaluating pilot’s requests has been sometimes found to be considerably slower under the visual (printed) datalink mode when compared to the traditional working mode (NRC, 1998). In a study by Wickens (1994), performance was best with the auditory-verbal request mode. It was recommended that the auditory channel should not be abandoned in the development of datalink.

The datalink example shows that the impact of automated tools on the communication tasks of the controller and the team is important. However, the changes will depend greatly on the way the way the tools will be implemented and on the way teamwork will evolve in order to support these changes.

CAST (1998b) refers to studies that considered the risk that with the introduction of datalink R/T communication skills cannot be maintained. Some believe that some R/T should be preserved in order to partially negate this. To the question: ‘How will interpersonal skills and teamwork skills be affected by silent coordination and datalink?’, the answer was that although electronic coordination may replace verbal coordination for the most part, it is likely that controllers will discuss options orally – especially if they are in close proximity with their colleagues in the operations room. Routine use of datalink for communication with pilots is likely to cause a reduction in the ‘professional relationships’ established between controllers and pilots. It is not clear what the overall effect of these changes will be. Which team tasks of Table 1 will be affected by datalink, will generally depend on the specific datalink application.

3.2.2 Automation in flight Information systems

Example application: Electronic Flight Strips (EFSs)

The flight plan must be acquired by the system, processed and associated with radar information. Already some decades ago, activities started to give some degree of automation to flight planning and flight progress strips, because it was recognized that manual processing of flight data within and between units is one of the main sources of workload for the controllers.

In the EUROCONTROL ODID III Study in the early nineties the performance of two controllers working side-by-side was compared with their performance when separated and communicating only through their displays and communication link. At the outset the electronic strip display contained too much data and there was no visual indication for the completed tasks and those that were still outstanding. The Planning Controller (PC), who had no PVD available, frequently needed to move over to the display of the TC. In a more advanced configuration, the planner also had a PVD and the amount of permanently visible flight data was minimised.

The EUROCONTROL ODID IV Study was a strip-less environment. All important flight data was contained into the Track Data Blocks (TDBs), which were connected to each aircraft plot on the PVD. These TDBs allowed efficient mouse-based data entry (updates) Visual indications of outstanding tasks were provided in the PVD, mostly through colour coding of the TDB or fields thereof. The ODID IV system also contained tools for strategic long-range planning (see Section 3.2.4). The HMI of ODID IV was an important basis for the aforementioned PHARE PD/1 system (EUROCONTROL, 1997).

NRC (1998) concludes that although controllers still work in teams, and their supervisors are called team supervisors, the en-route and TMA flight information processing systems permit controllers to work more independently than they did before the introduction of these systems, with team members paying less attention to each other. In that respect, the trend is probably the same in Europe, due to the introduction of silent coordination, automatic hand-offs, flight-data presented in TDBs on the PVD, etc.

3.2.3 Automation in immediate conflict avoidance

Example application: Converging Runway Display Aid (CRDA)

When airports use converging runways, i.e. when the final approach paths, or the extended centrelines, of the two active runways cross each other, Approach Control can choose to implement a software tool called the Converging Runway Display Aid (CRDA). CRDA is a tool that helps controllers in sequencing the traffic flows that arrive at the converging runways.

When traffic load calls for the use of two converging runways, then the approaching aircraft on the two different descent paths need to have a relative sequential spacing according to some separation minimum. If this relative sequential spacing is not maintained, separation criteria would be violated and aircraft would collide at the point where the final approach paths cross each other. Thus the critical stage is at the point at which the two approach paths come together; the two arrival sequences must be staggered at this point.

Also, if the extended centrelines of the two active runways cross each other (rather than the approach paths), then a specific relative sequential spacing should be maintained in order to avoid loss of separation in case of a double missed approach (i.e. when two approaching aircraft must go-around).

The separation task can be difficult, e.g. when weight differences (heavy, medium and light aircraft) give rise to speed variations between aircraft. Without the CRDA the controller must use range-rings on the radar display to determine the relative sequential spacing between aircraft in different approach paths, say A and B. However, the CRDA shows each aircraft on approach path A also as a ‘ghost image’ on approach path B, and vice versa (specific implementation depends on the airfield). Each ghost image is created by rotating the position of each aircraft on path A around the intersection point over an angle that is equal to the angle between the two converging runways.

As a result, the in-trail separation of the aircraft on both paths can be directly compared against a criterion and controlled.

The system has been evaluated (among many other airports in the US) at Schiphol airport in Amsterdam, at the National Aerospace Laboratory using the NARSIM air traffic control simulator. A potential obstacle for implementation was related to skill decay, since CRDA could only be used under specific (low) visibility conditions. Actual practice with CRDA could be rare in periods with good visibility, such that periodic refresher training would be needed to exercise the skills and procedures associated with CRDA.

Under high traffic load, two controllers would be needed for each of the two runways, i.e. each would build a sequence for one runway. With CRDA installed, each controller would see the aircraft under the other’s control as ghost images. It is conceivable that this circumstance could reduce the need for communication and explicit coordination (NRC, 1998). Hence teamwork could be influenced with this automated tool for immediate conflict avoidance.

3.2.4 Automation in strategic long-range planning

Example application: Highly Interactive Problem Solver (HIPS)

In the aforementioned EUROCONTROL PHARE Program a set of integrated automatic tools was developed with which controllers could determine conflict-free trajectories and negotiate these (semi-)automatically with the aircraft. Principal components of the system included an experimental flight management system in combination with bi-directional datalink capability and controller tools such as the ‘trajectory predictor’ and the ‘conflict probe’. At the MMI level, these advanced tools are embodied in the Highly Interactive Problem Solver (HIPS) (Meckiff & Gibbs, 1994), which allows the controller to view, edit, negotiate and approve trajectories through direct manipulation (mostly drag and drop) of the trajectories.

While the HIPS is activated for a certain aircraft, its three displays will open. In the bottom left corner a ‘route display’ is presented. This display is used to identify problems in the horizontal plane of the subject aircraft and to implement heading changes in the planned trajectory of the aircraft. It is called the Horizontal Problem Solver (HPS). Similarly, the ‘speed display’ (top right corner) and the ‘altitude display’ (bottom right corner) are called the Speed Problem Solver and the Altitude Problem Solver, respectively. A solution is found through manipulating the trajectory such that it does not cross a ‘no-go’ zone, i.e. the red (conflict) and yellow (conflict risk) areas on the HIPS displays. The trajectory support tool is used to validate, register, propose or reset a trajectory. Note that the size of the PVD must be reduced to allow space for the HIPS displays (EUROCONTROL, 1997).

Human Factors considerations appear to have been paramount at all stages of the HIPS development (Jorna, 1997). Subject workload data collected during the PD/1 trials (involving 32 controllers from eight different ATC organisations, which received one week of training and one week of experimental trials) suggested a better distribution of workload between tactical and planner, while in the baseline condition (without the PHARE tools) the tactical reported a higher workload than the planner. Generally, controllers were particularly excited and positive about working with the HIPS.

With respect to teamwork it was noticed that the design of HIPS was guarded for ‘team mode confusion’, i.e. confusion of who is responsible for what task at what time. The planner controller could work with the HIPS in either look-ahead or real-time mode, but not communicate with the aircraft, whereas the TC’s HIPS display operated in real time only (NRC, 1998).

Because the HIPS is designed to support long term strategic planning, it has the potential to shift more control from tactical to the planner. In fact, one can think of the situation in which the tactical only implements the long-term plan that has been sorted out by the planner using the HIPS, thereby ensuring conflict free flights with little need for active tactical control. Thus, while the team of planner and tactical as a whole does not necessarily suffer loss of skill in general, there may well be a relocation of specific skills from tactical to planner in such scenario. This leaves the team more vulnerable in issuing rapid tactical control commands to avoid conflicts when the automation (i.e. the HIPS) fails.

In subsequent PHARE trials, particularly PD/3 (EUROCONTROL, 1998), in which a new en-route multi-sector planning concept was introduced, more experience with the HIPS, which was then called PHARE Advanced Tools (PATs) Problem Solver, was gained.

3.2.5 Automation integration

Example application: Collaborative Decision-Making (CDM)

In the validation phase of the current project the revised SHAPE Teamwork Measure will be partially validated in a simulated aerodrome environment. This simulated environment will be created in the ‘Linking Existing ON-ground ARrival and Departure Operations (LEONARDO)’ European Union Project.

LEONARDO aims at the integration of airport traffic management systems: arrival and departure management and ground movement routing and planning. More specifically, one of the automation concepts that will be tested in LEONARDO is Collaborative Decision-Making (CDM).

Collaborative Decision-making (CDM) refers to a set of applications aimed at improving flight operations through the increased involvement of (1) airspace users, (2) ATM service providers, (3) airport operators and (4) other stakeholders in the process of air traffic management. Collaborative decision-making applies to all layers of decisions, from longer-term planning activities to real-time operations, and is based on the sharing of information about events, preferences and constraints. The set of applications that will be empirically tested in the LEONARDO CDM trials consists of:

- Departure Manager (DMAN), optimising the sequence of departures from one or more runways;
- Arrival Manager (AMAN);
- Gate Allocation System (GAS);
- Taxi-In Planner (TIP);
- Taxi-Out Planner (TOP).

The nature of the impact of CDM on teamwork will be measured using the refined SHAPE Teamwork Measure. The teamwork measures will be validated against other teamwork measures (based on eye movement patterns). The experimental validation plan will be described in Section 6.

3.3 Analysis of Team Tasks to Determine Skills, Knowledge and Attitudes

Prerequisite to the measurement of team performance is the team task analysis. A team task analysis provides information about team skills, knowledge and attitudes needed. The basis for such an analysis could be the task list of Table 1. The analysis could identify the cues, events, actions, coordination demands and communication flows needed for effective teamwork. It would enable the understanding of the nature of task interdependency in teamwork and to distinguish collective team tasks from individual tasks (Salas & Cannon-Bowers, 2000). Although we will use previous results from literature for the current purposes, the EATCHIP Human Resources Team (HRT) (1998a) recommends the following two methods for analysis of team tasks:

- analysis of team communication using protocol analysis,
- group flight-progress reconstruction.

Analysis of team communication

One category of Cognitive Task Analysis (CTA) methods outlines team communication. A close link between CTA and Team Resource Management (TRM) can be obtained by using a team communication method. In most cases, various types of protocol analyses have been used to study communication among team members. Socio-graphic methods may also be employed to graphically depict (1) direction, (2) frequency and (3) directness of communication. In addition, analyses of frequency, time pattern and pitch as non-verbal communication can be recorded and analysed using computerised techniques.

For recording team communication, use of the Information Flow Sheet is mentioned: The communication partners (e.g. pilots, planners, coordinators and adjacent sector controllers) and the frequency and direction of the communications are recorded.

Group flight-progress reconstruction

Flight Progress Reconstruction is a method to obtain an almost perfect recollection of the events in the observational period. As most of the relevant information is documented on the Flight Progress Strips (FPSs), they are collected and used as reconstruction tools for the analysis (either paper strips or printouts from electronic strips). The comments (reported Phase II of the investigations of the EATMP HRT [1999a]) indicated that, for the planner position, it was possible to use the flight progress reconstruction technique to make the decision-making and the working strategies of the PCs more transparent.

When analysing teamwork the HRT advises to collect (in addition to data resulting from the task analysis) organisational data (e.g. information on the structure of ATC unit) including size, workplaces and different positions, hierarchical structure, the shift system and team structure, the sectorisation, and the 'automated support tools'.

The next section will give an account of teamwork skills, knowledge and attitudes on the basis of literature which is based on these or similar methods for team task analysis.

3.4 Conclusions regarding Automation and Team Tasks

We considered the impact of automation on tasks. We started with a consideration of Levels Of Automation (LOAs). A one-dimensional scale for LOA is the 'Scale of Levels of Automation of Decision and Control Action' (Parasuraman *et al.*, 2000). NRC (1998) recommended that if a system is associated with high uncertainty and risk, then the automation should do no more than selecting and suggesting decision alternatives (such as alternative conflict-free trajectories). The controllers should then decide between the alternatives and execute the control actions.

Additional scales for LOA were suggested by NRC and EUROCONTROL, since it is clear that LOA cannot be fully captured by a one-dimensional scale. Other LOA scales try to capture the level of information automation.

In the design phase of ATC systems functions of the system (e.g. surveillance, communication, etc..) can be assigned to the machine (machine-functions), an individual operator (individual task) or a team of operators (team task). Scrutiny of task analysis documents identified 24 top-level team tasks (see Table 1), of which approximately two third are normal (routine) tasks and one third non-normal (emergency) tasks. For most of these team tasks some level of automation assistance is possible, and this may have been demonstrated in operational or research applications. However, some tasks involving complex judgements and other higher cognitive functions are almost impossible to automate. This is the case of the following high-level tasks:

- monitor fellow team members for performance, SA and workload;
- engage in the splitting and merging of functions, sectors or areas to cope with team task load;
- anticipate each other's reactions, capabilities and acceptance levels.

Particularly, tasks that need to be exercised in non-normal situations are rare and many non-normal situations simply cannot be anticipated, let alone automated. For example, the general team task 'manage a degraded system and ask for human support for tasks normally accomplished by a single controller' may be able to automate under very specific anticipated conditions, but must usually be resolved by the team.

4. TEAM SKILLS, KNOWLEDGE AND ATTITUDES

4.1 Team Skills

In the previous section we described the dimensions and levels of automation. The two main dimensions are 'decision and action selection' and 'information automation'. We also made a distinction between individual tasks and team tasks and these have been listed. Team tasks involve two or more controllers, together responsible for a sector or aerodrome area. These controllers have the common objective to control the flow of traffic in a safe and efficient fashion, which requires coordination and cooperation in the fulfilment of team tasks. In this section the knowledge, skills and attitudes required for teamwork are investigated. A skill is defined as a goal directed, well organised behaviour that is acquired to practice and performed with 'economy of effort' (e.g. with respect to expended time, accuracy and energy). Skill is built up gradually in the course of repeated training or other experience. Instead of skill, sometimes the terms ability or competency are used. Team skills require the integration of certain knowledge, social behaviour and attitudes such that the team is able to control the traffic. Good team skills allow the team to improve (optimise) their performance, and adapt it to the situation, such that the team is able to cope with both routine and non-normal tasks.

We will dwell upon the notion that 'coordination' and 'cooperation' are the two classes of skills needed in the fulfilment of team tasks. We will define coordination skills as a set of skills that is 'outcome-' or 'product-oriented'. The overall product of an air traffic control team is delivering air traffic services, or more generally stated, the safe and expedient flow of traffic. However, outcomes of team tasks can also be decisions, information updates, or actions. Team tasks in ATC (see Table 1) that require predominantly coordination skills are, for example:

- coordinate airspace reservations / non-standard flights,
- assess the traffic situation, seek information,
- provide an update on the traffic situation.

We will define cooperation skills, on the other hand, as a set of skills that is 'team process oriented'. These skills focus on the processes of teamwork, rather than the products of teamwork. ATC team tasks from Table 1 that require predominantly cooperation skills are, for example:

- monitor fellow team members for performance, situation awareness and workload;
- anticipate each other's reactions, capabilities and acceptance levels;
- in non-normal situations or emergencies, ask for human support for tasks that are normally accomplished by a single controller.

4.1.1 Team skills in applied literature

Many recent literature sources have tried to classify team skills. In order to provide an overview of these skills and in order to be able to classify the most important of these skills either in the class of coordination skills or the class of cooperation skills, we will deal with the applied literature (human factors / ergonomics, applied to teamwork in complex task settings). Team skills frequently mentioned in the literature are:

- <u>Team monitoring</u>

  mutual performance monitoring , mutual workload monitoring, predicting each others' behaviour (Hackman, 1990; Volpe *et al*., 1996).

- <u>Team flexibility</u>

  adaptability, adapting to novel and unpredictable situations, exhibiting flexibility (Prince & Salas, 1993).

- <u>Team leadership/followership</u>

  motivating team members, team initiative, exhibiting assertiveness, supporting behaviours (Smith-Jentsch *et al.*, 1996, 1998; Cannon-Bowers *et al.*, 1995; Salas & Cannon-Bowers, 2000).

- <u>Team coordination</u>

  - giving suggestions or criticisms, acceptance of suggestions or criticism, performing self-correction (McIntyre & Salas, 1995; Morgan *et al.*, 1986);

  - response coordination, coordination activities, resource distribution, timing, interpersonal coordination, team decision-making, shared situation awareness (Nieva *et al.*, 1978; Kleinman & Serfaty, 1989; Smith-Jentsch *et al.*, 1998).

The teamwork scale of Adult Literacy and Lifeskills (ALL) Survey (Smith, 2000) also seeks to assess the core skills associated with teamwork. To this end three primary skills required for effective teamwork - (1) group decision-making/planning, (2) adaptability/flexibility and (3) interpersonal relations – were proposed, each represented by distinct behaviour.

1. Group decision-making/planning refers to the ability to identify problems and gather, evaluate, share and link information.

2. Adaptability/flexibility implies using a variety of task-relevant strategies, providing assistance, adjusting to task reallocation and accepting feedback.

3. Interpersonal relations are apparent from supporting team decisions, sharing work, helping others and seeking mutually agreeable solutions.

Smith (2000) further argues that communication skills, including providing complete and concise information, listening effectively, and asking questions – underlie the other three skills and serve as a bridge among them. The NOTECHS Consortium (Avermaete & Kruijsen, 1998) takes a similar view of communication as a skill: *... communication is not so much a skill category as well as a means to be able to perform on each of the other categories*.

Focusing on Non-Technical Skills (NTS) rather than merely on teamwork skills, NOTECHS (1998) distinguishes four categories for the NTS of commercial flight crew:

1. Cooperation.
2. Leadership and managerial skills.
3. Situation awareness.
4. Decision-making.

Each category is subdivided in three or four elements. Each element is coupled to one or more behaviours. For example, situation awareness category has an element 'environmental awareness' which is coupled with the behaviours 'collects information about the environment', 'contacts outside resources when necessary', 'shares information about the environment with others'.

In ATM core skills (not necessarily team skills) of air traffic controllers are (Pearn & Kandola, 1983; Eissfeldt, 1988, 1991, 2002; Cox, 1994):

- the ability to process (visual and acoustic) information simultaneously from multiple sources (selective attention) to build up and maintain the traffic picture;
- the ability to absorb new information while making decisions (quickly and accurately);
- the ability to project forward (i.e. planning prospective actions) on the basis of current information;
- the ability/flexibility to constantly adjust the whole picture and modify plans.

CAST (1998b) predicts the following 'key' skill changes of controllers in the future (i.e. ca. 2015):

- The controller is likely to become more passive (i.e. a more supervisory role) – having less direct input to the control activities. As a consequence, the controller will get less practice in the use of various skills, or must practice these in other ways.
- The controller may not be so aware of the global context of his/her sector – focusing instead on small areas.
- The controller must have the ability to cope with non-normal situations or emergencies that cannot be solved by the automation. Generally, problem

solving will be called for less, but when it is called for, it will be harder and in these cases the controller must be able to intervene appropriately when it becomes necessary.

- The controller must have more insight in the capabilities, the behaviour and the function of the (automated) facilities that are in place to support him.

The information that formed the basis for the skill analysis was derived from a cognitive task analysis with prototype systems.

An earlier report of CAST (1998a) concluded: In the future, collaborative decision-making takes place by a permanent dialogue between all partners and systems involved. There will be more cooperation with computers and other operators, both in the air as well as on the ground. The dynamic nature of air-ground responsibility would place greater emphasis on information sharing and handling air-ground relationships.

According to VINTHEC (2003) one of the important constructs that facilitate teamwork is Team Situation Awareness (TSA). Building and maintaining TSA is a skill that teams as well as individual operators need to master. Insight and training are the main building blocks for operators when they are learning to master these skills. Being able to measure when, or if, this skill is mastered is useful for a number of reasons. In order to measure a skill like ‘building and maintaining TSA’ the skill is split up in a number of behaviours which, in their turn, can be split up in separate measures.

The behaviour that may reveal aspects of TSA comprises (VINTHEC, 2003):

- <u>Communication</u>

  Not just the content of the words exchanged between individuals, but also the non verbal aspects like intonation, facial expressions, pointing, etc.

- <u>Paying attention to…</u>

  One of the key issues of TSA is paying attention to relevant aspects of the environment. What relevant aspects are is dependent on the situation at that moment. Information may come via (boxes on) displays, via telephone, verbal or non-verbal expressed by other operators.

- <u>Planning ahead / thinking about the future</u>

  The teams as well as the individuals need to plan their steps in order to remain ahead, and in control, of the situation. Therefore they need to actively process information on their own and in the team. This processing is a difficult to measure type of behaviour, though aspects of it may be measured. These express themselves in the content of communication, but also in the type of information that operators look at / read.

Currently, SHAPE distinguishes the following top-level team **processes**:

1. Communication.
2. Cooperation (facilitation).
3. Coordination.
4. Situational Awareness (SA).
5. Leadership.

Team **characteristics** that distinguish teams from small groups include the following (Dwyer, 1984; Modrick, 1986; Morgan *et al.* 1986; Salas & Cannon-Bowers, 1997):

1. Multiple sources of information.
2. Task interdependencies.
3. Coordination among members.
4. Common and valued goals.
5. Specialised member roles and responsibilities.
6. Task-relevant knowledge.
7. Intensive communication.
8. Adaptive strategies to help respond to change.

According to Masson and Paries (1998), teamwork in ATM teams is characterised by six **features**:

1. Team structure.
2. Leadership and command.
3. Communication.
4. Decision-making.
5. Situation Awareness (SA).
6. Cultural aspect.

Masson and Paries stress that a long lasting ATM team functions as a ‘clan’, including self-organisation, accurately defined roles and statuses, peer pressure, strong team identity, team culture, etc.

Chung, O’Neil and Herl (1999) investigated measures of teamwork skills in computer supported cooperative work. They studied the nature of the interactions between team members as they jointly constructed a knowledge map. The interactions between team members consisted of sending other team members pre-defined messages. Each message was categorised as belonging to one of the following **team processes**:

1. Adaptability.
2. Communication.
3. Coordination.
4. Decision-making.
5. Interpersonal.
6. Leadership.

Paris *et al.* (2000) state that in addition to developing theories and models, team researchers have struggled to identify critical skills or traits that enable teams to deal with relevant task information, i.e. **skills or traits** that enable teams to:

1. Coordinate.

2. Communicate.
3. Strategize.
4. Adapt.
5. Synchronise.

On the basis of extensive literature review Dickinson and MacIntyre (1997) have defined seven main **components** of teamwork:

1. Communication.
2. Team orientation.
3. Team leadership.
4. Monitoring.
5. Feedback.
6. Backup behaviour.
7. Coordination.

For measurement purposes in a naval warfare context, Dwyer *et al.* (1997) distinguished four teamwork **dimensions**:

1. Situation assessment.
2. Communication.
3. Supporting behaviour.
4. Leadership.

Morgan *et al.* (1986) describe the developmental phases through which teams typically evolve (according to the 'team evaluation and maturation model') on seven **dimensions**:

1. Communication.
2. Cooperation.
3. Morale.
4. Adaptability.
5. Coordination.
6. Acceptance of suggestions.
7. Giving suggestions.

Nieva *et al.* (1978) identify **antecedent factors** that impact upon team performance:

1. Member resources.
2. Team characteristics.
3. Task characteristics.
4. Task demands.

The pre-study refers to the MATT Methodology, which, at the highest level, distinguishes four team **processes**:

1. Communication.
2. Coordination.
3. Adaptive behaviours.
4. Backup behaviours.

Entin and Entin (2001) developed a teamwork measure comprising of fifteen items that measured six **dimensions** of teamwork:

1. Communication.
2. Monitoring.
3. Feedback.
4. Backup.
5. Coordination.
6. Team orientation.

The Line Operation Safety Audit (LOSA) System, which, like NOTECHS, is meant as an evaluation system to be used during airline crew training/checking in a flight simulator environment distinguishes only three **skill** categories:

1. Planning.
2. Execute.
3. Review or modify plans.

4.2 Teamwork Knowledge

Information on the specific knowledge of team members that enables them to function in an ATC team can be found in various EUROCONTROL EATCHIP/EATMP/EATM documents and lesson material. Sources are, among others:

- the EUROCONTROL TRM course material;
- the human factors modules (see EATCHIP, 1996a, b, 1997a, b, 1998b, c; EATMP, 1999b, c);
- the 'Team Resource Management Test and Evaluation' document (EATMP, 1999d);
- the 'Proceedings of the Second EUROCONTROL Human Factors Workshop - Teamwork in Air Traffic Services' (EATCHIP, 1998d).

Examples of such specific knowledge items that are important for teamwork in ATM are extracted from these sources and summarised below.

<u>Teamwork in ATM - general</u>

- What is teamwork in ATM?
- Who are your teamwork partners?
- Team identity
- What makes a good team?
- The positive and negative effects of a team structure

- How to cope with team pressure?
- How to create and retain a good teamwork atmosphere?

<u>Team roles</u>

- Understanding your roles in the Unit/Centre
- Attitudes towards authority
- Leadership and followership
- Dealing with submissiveness, assertiveness and aggressiveness
- Supervision and expertise
- The effect of errors on authority

<u>Communication</u>

- What is good communication?
- Identify the functions of communication
- Identify how communication is performed within teams and how it can effect safety
- Develop strategies on how to communicate effectively
- Develop strategies on how to intervene effectively in a typical ATM-related situation
- Develop ways to give and receive feedback and constructive criticism
- The ‘law of silence’

<u>Situational Awareness (SA)</u>

- Retaining awareness of own as well as other’s situation
- Identify the symptoms of loss of individual and team SA
- Develop appropriate strategies on how to prevent the loss of (team) SA
- Identify factors that may have a positive or negative influence on (team) SA

Decision-making

- Understanding decision-making strategies and individual differences within these
- Appreciate the concepts of shared problem models and the use of resource management skills in team decision-making

Working and personal style

- Appreciating different controlling styles and appreciating that both leaders and followers can support
- Cultural differences (between units, organisations, regions and nations)

Stress

- How do you recognise your own as well as someone else's stressors?
- Explain how stress effects teamwork
- Develop skills to recognise and cope with stress situations in teams

Obviously, there is much more to say about the knowledge underlying successful teamwork, and this knowledge is acquired. Literature on the topic is abundant. However, there is an overlap between the knowledge underlying team skills and the skills of the team. Knowledge that has been effectively acquired will enable the development of skills through practice. For example:

- knowledge about strategies on how to communicate effectively may enable the skills to communicate effectively;
- knowledge of the effect of stress on teamwork, allows the team to recognise the symptoms and develop teamwork skills for coping with stress.

Therefore, in further development of a framework for teamwork skills, knowledge and attitudes, we will not attempt to distinguish specific types of teamwork knowledge, but assume that a certain knowledge base is a pre-requisite for the development of skills.

4.3 Teamwork Attitudes

Team attitudes are defined as an internal state that influences a team member's choices or decisions to act in a particular way (Cannon-Bowers *et al.*, 1995). Attitudes toward teamwork can have a significant effect on how teamwork scales are actually put into practice (Smith, 2000). Positive attitudes toward teamwork and an attraction to being part of a team ('collective orientation') have been found to enhance team processes and team performance. Also, attitudes towards teamwork depend on culture (JAR-TEL, 2001). Therefore, for some applications these attitudes should be addressed as part of the teamwork measure.

A significant body of work exists on the assessment of attitudes toward teamwork in aviation (see for instance Helmreich *et al.*, 1986), mostly focussing on commercial pilot attitudes toward teamwork in the cockpit. Some important attitudes found in the general literature are:

- team spirit,
- team morale,
- belief in the importance of teamwork,
- team cohesion,
- shared vision,
- mutual trust,
- collective orientation.

(Nieva *et al*., 1978; Driskell & Salas, 1992; Mullen & Copper, 1994; Gregorich *et al.*, 1990; Ruffell-Smith, 1979; Morgan *et al.*, 1986; Cannon-Bowers *et al.*, 1995; Salas & Cannon-Bowers, 2000).

For the measurement of attitudes towards teamwork some form of Likert scaling is usually favoured. Likert-type scales include a series of positive and negative statements about teamwork, and respondents endorse one of a series of graded response options (e.g. strongly disagree, slightly disagree, neutral, slightly agree, strongly agree, respectively). Points are allocated to each response option (e.g. one-five). The sum of these values represents attitude strength.

Motamedzade *et al.* (2002) recently designed a questionnaire that measured (among other things) attitudes towards teamwork in an industrial environment. Questions, using a five-point scale, referred to topics such as:

- size adequacy of the work team;
- ability of team members including technical expertise, problem solving and decision-making;
- personal and interpersonal skills;
- clarity of roles in team;
- having a vision and commitment to it;
- establishment of goals at the team level;
- leadership and structure at the team level;
- accountability of team members at both individual and team level;
- evaluation and reward system in team;

- developing high mutual trust in the team;
- continuous learning process as one of the main functions of team;
- willingness to protect each other and maintain identity of the team;
- conflict resolution in team;
- changing attitude towards the organisation;
- behavioural changes at both personal and team level;
- sense of ownership toward work done by team members;

Sherman and Helmreich (1995) used the Flight Management Attitude Questionnaire (FMAQ, developed for NASA by Helmreich and his colleagues) to measure the attitudes of pilots towards automation. They report that three main factors affect this attitude:

- National culture,
- time spent flying automated aircraft,
- seniority of crew member.

The earlier-mentioned questionnaire (see 2.3) by Flin *et al.* (2003) measures attitudes towards:

- leadership,
- communication,
- teamwork,
- stress,
- fatigue,
- work values,
- human error,
- organisational climate.

EATMP (1999a) reports on the Air Traffic Control Safety Questionnaire (ATCSQ), which was developed to enable the evaluation of the EUROCONTROL Team Resource Management Programme. The questionnaire includes a section that measures the attitudes of controllers. The results, based on a response sample of 29 controllers - which was considered rather small, reveal that the course changes attitudes in favour of better and more cooperative teamwork and more sympathetic team roles.

In one section of the questionnaire (consisting of forty items) responses are given on a five-point Likert scale (from ‘strongly disagree’ to ‘strongly agree’). Items relating both to attitudes towards automation and attitudes towards teamwork among others include:

- ‘automation reduces the requirement for team members to monitor the traffic situation closely’;

- ‘team members share responsibility for prioritising activities in high workload situations’;
- ‘asking for assistance makes one appear incompetent’;
- ‘to resolve conflicts controllers should openly discuss their strategies with each other’;
- ‘during periods of low work activity I would rather relax than keep busy with small tasks’;
- ‘increased automation reduces the need for team communication’.

A matter related to attitudes was addressed in CAST (1998b). In this document it was stated that if automation is widely implemented in ATC, new controllers will find it increasingly difficult to develop an understanding of the ‘professional norms, standards, and ethos’ in ATC that are associated with the current unquestionable dependency on humans in the ATC system. According to the CAST participants, this potential problem could be circumvented by ensuring that tasks in future ATC systems in which automation plays a significant role, require sufficient interaction between controllers to foster the perpetuation of professional skills.

4.4 Proposed Theoretical Framework for SHAPE

Since teamwork is a complex issue by itself, it was decide to postulate a model, the ‘Skills, Knowledge and Attitudes for TEamwork (SKATE)’ Model, that is rather ‘parsimonious’.

The sufficiency of this model for defining teamwork measures will be investigated in the remainder of the project. With ‘parsimonious’ we mean here that the model is based on minimal assumptions, definitions, and taxonomies. We will elaborate the SKATE Model, only when the validation experiment reveals influences by automation (which is the whole purpose of SHAPE) that cannot be described or explained with the current version of the model.

The core teamwork skills identified from the previous analysis are:

- coordination,
- cooperation.

Coordination is defined as a set of teamwork skills that is ‘outcome’- or ‘product’-oriented. These skills (and their related measures) concentrate on measuring the product of teamwork, i.e. teamwork in a particular context, at a particular point in time. A product or outcome can be for instance a decision, an information update or an action.

The concept of a ‘teamwork product’ does not expose how the decision was reached, how the information was obtained or how the team came to an action. This question ‘how?’ refers to the teamwork process.

Cooperation is defined as a set of skills that is team process oriented. Examples of skills that fall under the heading of cooperation are the skills necessary to unload a team member when s/he becomes too busy and to ensure that fellow team members understand your plan.

Clearly, in order to measure teamwork or to understand how teamwork develops it is necessary to understand how air traffic controllers perceive each other, e.g. each others' performance, each others' expertise, each others' workload, and the team constraints that play a role in processing that information. This is also captured by the term cooperation.

A further subdivision was made between leadership-type skills and followership-type skills. Some of the leadership-type skills are 'coordinative', i.e. focusing on the product of teamwork, while others are cooperative, i.e. with a focus on the teamwork process. The same distinction has been made for the follower-type skills. This corresponds with the notion that team members need both leadership and skills in order to make the team functioning.

Finally, the SKATE Model emphasises the attitudes towards teamwork. Some general attitudes towards teamwork are distinguished. These general attitudes have a positive effect on teamwork in general. Again, a further subdivision has been made in 'cooperative attitudes', i.e. influencing the teamwork process and 'coordinated attitudes', attitudes that on the one hand influence the teamwork product, but also require coordination before these attitudes can be instilled. For example, the attitude 'shared vision' is only possible when team members coordinate their individual visions.

Table 2:The SKATE Model

<table>
<tr><th colspan="4">SKATE
Skills, Knowledge and Attitudes for TEamwork</th></tr>
<tr><th colspan="4">Core Skills</th></tr>
<tr><td colspan="2">Team co-operation
(team process oriented)
• Monitor/assess each other's performance;
• Monitor/assess each other's workload;
• Monitor/assess each other's SA;
• Predict each others behaviour;
• Prevent future overloading of the team;
• Back-up others to prevent overloading;
• Adjust to each other's working style.</td><td colspan="2">Team co-ordination
(outcome oriented)
• Ensure which tasks are:
✓ Entirely own responsibility;
✓ Shared with other team members;
✓ Performed by the automated system.
• Assess the traffic situation / seek information;
• Provide updates on the traffic situation;
• Co-ordinate with other team members;
• Make team decisions;
• Select course of actions;
• Synchronise team actions.</td></tr>
<tr><td>Leader type qualities</td><td>Follower type qualities</td><td>Leader type qualities</td><td>Follower type qualities</td></tr>
<tr><td>• Redistribute workload when required;
• Take initiative;
• Motivate the team;
• Tactful alert other members to mistakes;
• Suggest ways to find and remedy errors.</td><td>• Exhibit assertiveness;
• Ask for assistance when necessary;
• Help others with difficult tasks or problems.</td><td>• Focus the team on its tasks;
• Form and disseminate plans;
• Assign tasks;
• Give orders and directives;
• Take command.</td><td>• Accept suggestions or criticisms;
• Perform self corrections;
• Provide an input or response when asked.</td></tr>
<tr><th colspan="4">Attitudes</th></tr>
<tr><td colspan="4">General attitudes towards teamwork:
• Belief in the importance of team work.
• Belief in the continuous learning process as one of the main functions of the team.</td></tr>
<tr><td colspan="2">Co-operative attitudes</td><td colspan="2">Co-ordinated attitudes</td></tr>
<tr><td colspan="2">• Team spirit, team morale, team cohesion;
• Willingness to maintain identity of the team.</td><td colspan="2">• Shared vision;
• Mutual trust in the team.</td></tr>
</table>

5. MEASURING TEAMWORK SKILLS, KNOWLEDGE AND ATTITUDES

5.1 Introduction

Previously, human performance investigations characteristically focused on the performance of individuals and did not address the measurements of teamwork skills (Baker & Salas, 1992). Dyer (1984) already noted that the lack of teamwork measures presented serious problems for conducting team-oriented investigations because investigations that involve measurements are often judged by the quality of the measurement instrument employed. To measure teamwork effectively, the chosen instrument should have a theoretical foundation, psychometrically sound, and a practically useful indicator of teamwork. Team assessment tools should (Paris *et al.*, 2000):

1. Identify processes linked to key team products.

2. Distinguish between individual and team level deficiencies.

3. Describe interactions among team members so as to capture the moment-to-moment changes that occur.

4. Produce assessments that can be used to deliver specific performance feedback.

5. Produce evaluations that are reliable and defensible.

6. Support operational use.

Metrics that fulfil these criteria will measure, detect and repair major skill deficiencies.

5.2 Outcome and Process Measures

5.2.1 Outcome measures

In ATM trials, outcome measures typically rely on opinion of Subject Matter Experts (SMEs) in combination with computerised performance recording. ATM outcome measures may include ratings of proficiency with respect to:

- speed (timeliness) of team response (e.g. search time, reading time, time to task completion, response time between conflict detection and clearance);

- accuracy/errors (e.g. based on aircraft spacing);

- number and type of omitted or incomplete tasks;

- frequency counts (e.g. frequency of tool usage, number of altitude transitions, number of aircraft accepted into sector, number of button pushes, number of hand-offs, and number of separation violations);
- duration/rate counts (e.g. Radiotelephony [R/T] average call duration, R/T usage rate);
- measures of knowledge.

5.2.2 Process measures

Although ATM outcomes reveal the extent to which automation design objectives have been met, consideration of *how* the task was accomplished (i.e. team process measurement) is important for diagnosing performance problems that may sooner or later obstruct or prevent desired ATM outcomes. Because unsafe team processes may occasionally result in successful ATM outcomes, it becomes necessary to measure both team processes and ATM outcomes if one is to ensure consistently effective performance.

Process measures might include:

- analysis of team communications or information flow,
- analysis of interactions or strategies employed by team members,
- interviews/debriefs,
- field observation techniques,
- task analysis,
- modelling or computerised simulation of teams,
- mathematical indices of team processes.

5.3 Measures Obtained from Observers, Participants and the Simulation

5.3.1 Measures obtained from observers and participants

Observer-based measures are obtained from knowledgeable, trained observers who observe team performance and processes during a scenario run (Entin & Entin, 2001). The ratings provided by observers can relate to individual members of the team, to sub-components of the team, or to the team as whole.

Participant-based measures can focus on self, other team members (individually or collectively), and team as a whole. In most cases, participant-based measures are typically obtained at the end of a scenario, but in some cases it is possible to halt the simulation at specified times to obtain interim assessments. For example, for measuring Team Situation Awareness (TSA), Hanson *et al.* (2003) describes both an ‘ecological momentary assessment’ method that uses a simulation freeze procedure and an ‘auto-confrontation’ method that relies on a post-session review to evaluate TSA.

Table 3 (taken from Entin & Entin, 2001) provides a high-level view of observer- and participant-based team measures that have been used in team research projects. The four columns in the table show respectively the name of the measure, the level at which it is captured (individual or team), the source of information (observer or participant) and an explanation of the measure. Entin and Entin stress that these measures can be obtained in virtually any simulator-based exercise. The particular measures that are collected in any given situation will depend both upon the factors incorporated into the scenario ant the goals of the experiment in which the measures are collected.

Table 3: Overview of observer- and participant-based measures (Entin & Entin, 2001)

Measure	Level (focus) of observation	Source of data	Description
Performance Outcome	Team	Observer	Ratings of overall team performance of aspects thereof.
Teamwork	Team	Observer	Quality of (dimensions of) teamwork processes (e.g. communication, coordination, etc.).
Team processes and dynamics	Individual, team	Participant	Enumeration of unobservable individual and team factors underlying team processes derived from scenario-based structured interviews.
Verbal communication	Individual	Observer	Records of type, sender and recipient, type and time of communications.
Workload	Individual, team	Participant	Assessment of individual workload for self and others; global assessment of team workload.
Mutual mental model congruence	Individual, team	Participant	Assessment of the congruence of models that team members hold of one another.
Organisational awareness	Individual	Participant	Assessment of the accuracy or congruence of team members' situational and mutual mental models.
Scenario and tools evaluation	Individual, team	Participant	Ratings of aspects of scenario including level of difficulty, complexity, uncertainty, ambiguity for self, others and/or team as a whole.
Attitude/climate evaluation	Individual, team	Participant	Ratings of attitudes, feelings, and opinions pertaining to selected issues or topics.

Measuring team communication

When communications among the team members are captured at a detailed semantic level, it is very difficult and time consuming to develop meaningful, quantitatively based measures to describe the nature of the communications. On the other extreme, simple frequency counts of communications, though straightforward, do not provide a meaningful window into team processes. In an approach described by Entin and Entin (2001), verbal communications among the members of the team are captured by observers at an intermediate level of detail that incorporates both semantic and quantitative aspects of the communication stream. Using the technique described by Entin and Entin, the observers listen to the communications in real time during a simulation scenario. The observers use a specially designed form to code the source, the recipient, the time (if a hand-held computer is used) and the type of verbal communications among the team members. Types of communication are divided in three basic categories: transfers, requests and acknowledgements. Both transfers and requests can, in turn, be classified as requests for information, action or coordination.

Measuring team workload

Team workload measures have been used extensively in various domains and have been found to be negatively related to measures of team performance. The Task Load Index (NASA-TLX; Hart & Staveland, 1988) provides an assessment of individual team members' workload. The TLX can be extended to capture team as well as individual workload (Entin, Serfaty & Kerrigan, 1998, cited in Entin & Entin, 2001). The latter authors used a revised workload measure. In the first part, participants reported their own workload in terms the traditional items comprising the TLX. In the second part of the questionnaire, each participant provides an estimate of the overall workload experienced by each of the other team members. In the third part of the questionnaire each participant responds to the TLX items, but this time for the team as a whole (not just for themselves).

A different approach to team workload has been used in determining the workload in a new cockpit that utilises a new team concept (Bohnen, 2003). In this case a combination of NASA-TLX and Cooper-Harper ratings have been used. According to this approach, the workload of a team member consists of four components (mental effort, time constraint, stress and physiological effort). After a simulated scenario segment, ratings and scores are obtained in three steps:

- Each team member gives a rating (from low to high workload) for each component, using TLX and the adapted Cooper-Harper Test.
- In addition each team member gives a unique weight (one to six) to each of the possible six combinations of two components, according to the weight experienced by this combination of components.

- In the third part of the questionnaire, each team member assigns four uncertainty scores to the rating for each component Each uncertainty score is related to an uncertainty source such as an uncertainty in the workload rating due to a non-representative scenario. The four uncertainty sources are (1) 'scenario', (2) 'simulator', (3) 'knowledge/training' and (4) 'system functions'. The team member judges whether he possibly made a rating error in the workload component as a result of an uncertainty source. Hence, four qualitative uncertainty scores per component are obtained (and a total of sixteen uncertainty scores for the four components that have been rated for the scenario).

- An individual workload value (midpoint), plus an uncertainty range around this value (defined by a lower and upper bound), can than be calculated.

- To obtain a workload score and uncertainty range for the whole team, the *maximum* value of midpoint, lower and upper bound is determined from all team members in the flight. Thus, the team member that experiences the highest workload is considered to be representative for the overall workload of the team (which seems fair for complex tasks with high interdependencies, for which the strength of the chain is in the 'weakest' link). To obtain an overall workload score for all teams participating in an experiment, the **harmonic** mean values of midpoint, lower and upper bound over all teams are determined.

- Finally, overall team workload functions (sawtooth-shaped functions with midpoint, lower bound and upper bound) obtained under different conditions (e.g. different levels of automation) can be compared.

<u>Measuring team situation awareness and mental models</u>

There is an important distinction between 'shared mental models' and 'team mental models'. The team mental model refers to the shared cognition in a team as a collectivity, but not shared among dyads of individuals. In the description of Shared mental models this is not allowed. (Klimoski & Mohammed, 1994). The words 'team mental models' do not allow for the notion of multiple levels of or sets of shared knowledge, rather they refer to the overall degree of similarity between the mental models of individual team members (Langan-Fox *et al.*, 2001).

In the previous section on team workload reference was made to a questionnaire devised by Entin and Entin (2001). In one part of the questionnaire each participant provides an estimate of the overall workload experienced by each of the other team members.

Apart from assessing team workload the questionnaire data can also be used to assess the congruence of the team's perception of workload across members of the team. The measure of congruence provides a window into the accuracy of the team members' mutual mental models of the team functioning, which can be considered part of the team situation awareness. The development of a shared mental model of interacting team members' tasks and abilities is considered important. There is evidence (see for example Orasanu, 1990) that well performing teams make use of such mutual mental models.

Another measure Entin and Entin (2001) refer to is a measure of organisational awareness. This measure provides a further window into the coherence of the team's situational and internal mental models. Organisational awareness relates to the knowing what task the other team members was doing during salient events within the scenario. In a questionnaire that measures this type of team situation awareness, a task categorisation is devised. The number of category matches between team members can be counted and a percentage agreement (congruence score) can be computed for each team.

Measuring team situation awareness with eye tracking and EPOG measurement

One of the measurement techniques that has proven its use in TSA assessment is EPOG recording and eye tracking. Without going into detail too far we'll very briefly explain the difference between EPOG and eye tracking. Eye tracking is just one of the subsystems of EPOG. Eye tracking measures the eye and can relate that information to video so that it becomes possible to actually see where someone is looking. EPOG combines information from a eye tracker with output of head tracker. This combination enables to digitally record at what predefined areas of interest a person had been looking. This type of data is less intuitive since there is no videotape that actually shows where a subject has focussed, but it is easier and faster to analyse since the data feed directly into excel or a statistical package. Of course a combination of both techniques is most ideal for research purposes.

Eye tracking and EPOG are measurement techniques enabling researchers to record where individual controllers or entire teams of operators were looking at. The relationship between pointing ones eyes in a certain direction, and actually looking at information, or even processing that information, is a tricky one. However, recording where someone focuses the eyes is still the closest that researchers can get to record where a person is paying attention to.

Because of the above, eye tracking is a popular technique to measure, amongst others, aspects of (T)SA. In the VINTHEC (I and II) studies (1999, 2000, 2003) eye tracking has proven its added value to questionnaires and other techniques when measuring (T)SA (VINTHEC, 2003). These studies took place in the cockpit, rather than in an ATC setting, where pilots (individuals as well as crews) where the subject of research. It was also demonstrated that there is a huge overlap with amongst domains such as anaesthesia, process control, nuclear power plant operations, maritime bridge operations and also ATC.

In all of these settings eye tracking can demonstrate whether:

- operators are looking at the relevant instruments at the right time;

- individual team members are looking at the same thing or that they are really sharing the work by each working on different aspects of the same control job;
- team members look at each others face (seeking confirmation in facial expression);
- team members look at each others hands (pointing, attraction each others' attention to something).

5.3.2 Measures obtained from the simulation

Research simulators often have the capability to log two kinds of data. One the one hand there are the inputs that the team or individual controllers gave to the simulator. On the other hand there are the results that were accomplished (i.e. the progress that the team/controllers have made in the simulated scenario).

In human factors research it is very common to compare and correlate several measures and to draw conclusions that are based upon the entire set of results. On the one hand there are measures that reveal information about how the subject perceives a situation. These can be questionnaire or rating scale based (either filled in by the subjects themselves or by external observers) but they can also be based upon psychophysiological measures or visual scanning behaviour.

However, the value of these teams'/controllers' perceptions is very much dependent on the teams'/controllers' performance (i.e. it is more tricky to compare reported workloads - no matter how they were measured - of two sets of teams/controllers if the team perceiving the lowest workload also performed below standard).

An efficient way to measure performance is to identify certain results that need to be accomplished by the team /controllers in the simulated scenario and to have the simulator log these measures of performance. These measures can both be, a number of inputs given to the system by the team/controllers, but also the progress that was made so far in the simulated scenario by the team/controllers. The 'objective' measures of performance that were logged by the simulator can afterwards be compared with the subjective rated assessment of performance (Entin & Entin, 2001), but also with other measures like workload, situation awareness, teamwork, trust, etc.

The following team-performance measures obtained from simulation have been used in the past in different research programs:

1. Aircraft spacing performance.
2. Amount of time no input was given to the system.
3. Amount of time that a certain information window was open / on screen.

4. Errors (all different kinds).
5. Event-based performance measures.
6. False alarm count/rate.
7. Frequency of use of different tools.
8. Job elements completed per unit time.
9. Miss rate / missed conflicts.
10. Number correct.
11. Number of aircraft accepted into sector.
12. Number of aircraft under control (at a specific (peak) moment, or average over the total duration of the run) per team or individual controller.
13. Number of altitude transitions.
14. Number of button pushes or mouse clicks.
15. Number of conflicts solved divided by the amount of conflicts encountered.
16. Number of errors.
17. Number of hand-offs.
18. Number of separated violations.
19. Number of speech acts / or duration of conversation (within team) needed to accomplish a certain goal.
20. Radiotelephony (R/T) average call duration.
21. R/T usage rate.
22. Rating of display clutter.
23. Response time between conflict detection and clearance.
24. Response time upon … numerous measurable facts.
25. RMS error.
26. Time to task completion.

5.4 Distinction between Quantitative and Qualitative Measures

Qualitative team measures can be ‘behaviourally anchored’ or ‘relatively low to high rating scale’. For example, team performance outcome measures capture the quality of the team’s performance on a scenario. In case of behaviourally anchored measure, an SME decomposes the scenario into its component team task and, for each task, specifies as the behavioural anchors the key behaviours that would indicate superior performance, adequate performance, and poor performance.

With behaviourally anchored scales, observers have fixed criteria against which to rate a team’s performance. In contrast, a relative low to high rating scale is sufficient for comparing one set of teams against another set, or for comparing team performance under different conditions (e.g. with and without certain automation features of tools), but does not allow for comparison to an absolute standard. The behaviourally anchored scale can be used to assess team performance on an absolute scale, a capability that is particularly important in situations such as team training exercise, where there is no comparison group or condition against which a team’s performance is being measured (Entin & Entin, 2001).

It is possible to assign numerical values to the ratings (e.g. poor performance=1, adequate performance=2 and superior performance=3), but statistical average of categorical data is always disputable because rating values do not have a numerical scale.

However, Entin and Entin (2001) use several qualitative observer ratings (validated in several earlier projects) on a behaviourally anchored seven-point scale to measure teamwork on six dimensions (communication, monitoring, feedback, backup, coordination and team orientation). They state:

> *The six dimensions of teamwork can be examined individually (for example, to look specifically at team backup behaviour, in subsets, or as a whole. The arithmetic average of the six ratings provides an overall measure of team process.*
>
> Entin & Entin (2001)

5.5 The SHAPE Teamwork Measure

5.5.1 Proposed revisions

The number of observers that was needed for handling Team Process Frequency (TPF) and Team Process Quality form (TPQ) is considerable, in particular when larger teams (more than two controllers) and analysis of inter-rater reliability is considered. Also, observer training may heavily burden an experimental budget. For example, Entin and Entin (2001) estimate that techniques to capture team verbal communication may require 25-30 hours of training to reach an acceptable level of proficiency. Therefore, we simplified the TPF and renamed it ‘observation form’. We used the approach of Entin and Entin to capture team verbal communication. The observation form, which includes an ‘observation matrix’ to record communication is included in Appendix C.

Analysis of the Observation form utilises anticipation ratios, e.g. the ratio of ‘information transfers’ to ‘information requests’. Such ratios have often proved more useful than individual rate measures for understanding team communication.

We revised the Team Process questionnaire (TP). The questionnaire focuses on automation-related attitudes towards teamwork. The FMAQ (Helmreich *et al.*, 1993), ORMAQ (Flin *et al.*, 2003) and ATCSQ (EATMP, 1999a) were used as examples. We are in the process of specifying the analysis as to quantify measured differences in automation-related attitudes as a result of exercises. The questionnaire is included in Appendix B.

We revised the Team Process Quality form (TPQ) and renamed it the ‘self-rating form’. We introduced ‘standardised’ self-assessment techniques (to be employed by controllers during or after the session), and modified for the current purposes, to capture aspects of team workload and team situation awareness. The ‘self-rating form’ is included in Appendix B.

5.5.2 Measuring team cooperation

Table 4 shows that each item falling under the heading ‘team cooperation’ of the postulated SKATE Model is covered by one of more items of the SHAPE Teamwork Measure.

Table 4: Measuring team cooperation

Team cooperation	**Covered with measure:**
• Monitor/assess each other's performance • Monitor/assess each other's workload • Monitor/assess each other's SA • Predict each others behaviour • Prevent future overloading of the team • Back up others to prevent overloading • Adjust to each other's working style	• Q – item 19, 26, 29, 30 • S (mutual ME ratings) • S (mutual SA ratings), Q – item 27 • S (mutual ME and SA ratings) • Q – item 30 • Q – item 6, EQ (ME and SA ratings for individual and team) • Q – item 8, 9
Cooperation – <u>Leader</u> type qualities	**Covered with measure:**
• Redistribute workload when required • Take initiative • Motivate the team • Tactful alert other members to mistakes • Suggest ways to find and remedy errors	• S (ME ratings for individual and team), Q – item 5 • Q – items 20, 25, 28 • Q – item 31 • Q – items 19, 20 • O (category 'suggest action or solution'), Q – item 28
Cooperation – <u>Follower</u> type qualities	**Covered with measure:**
• Exhibit assertiveness • Ask for assistance when necessary • Help others with difficult tasks or problems	• Q – items 6, 22, 28 • Q – items 6, CF (categories 'ask action' or 'ask information') • O (category: 'suggest action or solution')

ME = Mental Effort - O = Observation form - Q = Questionnaire - SA = Situational Awareness

5.5.3 Measuring team coordination

Table 5 shows that each item falling under the heading 'team coordination' of the postulated SKATE Model is covered by one of more items of the SHAPE Teamwork Measure.

Table 5: Measuring team coordination

Team coordination	Covered with measure:
• Ensure which tasks are:	
√ entirely own responsibility,	• Q - item 16, 18
√ shared with other team members,	• Q - item 1
√ performed by the automated system.	• Q - item 3, 4
• Assess the traffic situation / seek information	• Q - item 2
• Provide updates on the traffic situation	• O (category 'ask/give information')
• Perform procedural coordination	• Q - item 33
• Make team decisions	• O
• Select course of actions	• O
• Synchronise team actions	• Q item 29
Coordination – Leader type qualities	**Covered with measure:**
• Focus the team on its tasks	• O, Q – items 22, 23
• Form and disseminate plans	• O (category 'ask someone to perform action')
• Assign tasks	• Q – items 1-5.
• Give orders and directives	• O
• Take command	• Q – items 21, 32
Coordination – Follower type qualities	**Covered with measure:**
• Accept suggestions or criticisms	• O (category 'acknowledge/accept')
• Perform self-corrections	• Q – item 23
• Provide an input or response when asked	• O

O = Observation form - Q = Questionnaire

5.5.4 Measuring attitudes towards teamwork

Table 6 shows which items falling under the heading ‘attitudes’ of the adopted SKATE Model is covered by one of more items of the SHAPE Teamwork Measure.

Table 6: Measuring attitudes towards teamwork

Attitudes (General)	Covered with measure:
• Belief in the importance of team work • Belief in the continuous learning process as one of the main functions of the team	• Q – item 12 • Q – item 17
Cooperative attitudes	**Covered with measure:**
• Team spirit, team morale, team cohesion • Willingness to maintain identity of the team	• Q – item 11 • Not covered
Coordinated attitudes	**Covered with measure:**
• Shared vision • Mutual trust in the team	• Q – item 14 • Q – item 13

Q = Questionnaire

5.5.5 Data collection with the teamwork measure

5.5.5.1 Self-ratings

Detecting changes in team mental effort as a result of changes in automation level

After each simulation run the Mental Effort (ME) of a team member is rated in the following way:

- Each team member gives a rating for ME for him/herself and each other team member, using the Rating Scale Mental Effort (RSME, Zijlstra, 1993).
- Via the Uncertainty Sources Questionnaire (USQ), each crew member assigns four uncertainty scores to the ME ratings. Each uncertainty score is related to an uncertainty source, for example an uncertainty in the ME due to a non-representative traffic scenario. The four uncertainty sources are (1) ‘traffic scenario’, (2) ‘simulator’, (3) ‘knowledge/training’ and (4) ‘system functions’. The team member judges whether s/he rated mental

effort {‘correct’, ‘higher’, ‘lower’, ‘don’t know’} as a result of an uncertainty source. Hence, four qualitative uncertainty scores are obtained.

Calculating individual and team mental effort

For the purpose of determining ME, only the ME self-ratings will be used in combination with the ‘uncertainty ratings’ (see the USQ in Appendix B). To identify changes in ME multiple linear regression analysis ‘MANOVA’ or a different analysis technique can be used. Bohnen (2003) used fuzzy set theory to analyse ME self-ratings in combination with the USQ. However, here we only deal with regression analysis.

The regression model

The dependent variable is the RSME rating. The qualitative level of automation is coded in an independent dummy variable LOA (Level of Automation). For the reference system, LOA=0, for the CDM system: LOA=1. The uncertainty ratings {‘correct’, ‘lower’, ‘higher’, ‘don’t know’) can be used as a qualitative co-variates. This requires coding in two dummy variables, say HIGH and LOW, as follows:

RSME rating	**HIGH**	**LOW**
‘Correct’	0	0
‘Lower’	0	1
‘Higher’	1	0
‘Don’t know’	1	1

Each different source of uncertainty can be considered co-variate and thus requires coding, using a pair of such dummy variables. Also scenario-specific variables such as mode/runway can be considered co-variates and coded in a dummy variable RWY (independent use of runways: RWY=0, mixed mode: RWY=1). Finally, team functions (Runway Controller, Ground Controller, Departure/Start-up Controller) must be coded using two dummy variables, say FUNRWY and FUNGND, as follows:

RMSE rating	**FUNRWY**	**FUNGND**
Departure/Start-up	0	0
Ground Controller	0	1
Runway Controller	1	0

The full regression model, to be used for the data analysis of RSME data, then becomes:

$$RSME = b_0 + b_1 \cdot LOA + b_2 \cdot HIGH_1 + b_3 \cdot LOW_1 + b_4 \cdot HIGH_2 + b_5 \cdot LOW_2 + ... + b_6 \cdot RWY + b_7 \cdot FUNRWY + b_8 \cdot FUNGND + e$$

in which the coefficient beta-zero is the intercept and epsilon the model residual error. If the level of automation (LOA) has a significant effect on mental effort (as measured with the RSME rating) then the coefficient beta-one will be significantly different from zero.

Extra dummy variables $HIGH_X$ and LOW_X must be added for each uncertainty rating (related to realism of traffic scenario, system functions and simulator as well as the knowledge/training provided). Coefficients beta-seven and beta-eight, associated with variables *FUNRWY* and *FUNGND*, specify individual differences in Mental Effort. The basic hypothesis underlying this measure is that the introduction of automation has a significant influence on Team Mental Effort. This hypothesis is accepted when the variable *LOA* (representing 'Level of Automation') is associated with a level of significance less than 0.05. The coefficient beta-one specifies the sign and magnitude of this effect (when significant).

Calculating mutual mental model congruence

In addition to providing measures of the level of individual and team mental effort, we plan to use the questionnaire data in an innovative method (discussed in Entin & Entin, 2001) for assessing the congruence of the team's perception of mental effort and situation awareness across members of the team. The measure of congruence provided insight into the accuracy of the team members' mutual mental models of teamwork. It is hypothesised that the development of shared situational mental models of the task environment, the task itself, and of interacting team members' tasks and abilities are used to generate expectations about how other team members will behave. There is an increasing amount of research evidence showing that high-performing teams make use of mutual mental models.

We want to evaluate the accuracy with which team members can estimate the individual Mental Effort (ME) and Situation Awareness (SA) of each of their fellow team members. Therefore, we calculate congruence or deviation measures that reflect the difference between a team members self-reported ME or SA and the estimates of his or her workload made by every other team member. Entin and Entin (2001) compute the root mean square as follows:

1. The self-reported ME / SA for an individual is subtracted from each team member's estimate of the ME / SA for that individual.

2. These differences are squared, summed and averaged for the team.

3. The square root of the average is taken.

The mutual mental model congruence measure has been shown to have predictive validity in that it co-varies with performance outcome and teamwork (Entin, 1999). Appropriate statistics to test the differences in this root mean square for different levels of automation will be investigated.

5.5.5.2 *Analysis of data collected with the observation form*

It has in the past been observed that certain types of automation tend to limit severely the social context cues (visual feedback and status cues) that are available in face-to-face communication and to constrain the depth of discussion and analysis that is likely to occur. Moreover, the time needed for teams to make a decision sometimes increases. On the other hand, automated tools and computer-mediated interaction tends to mask status differences, resulting in greater participation by members and possibly better team coordination (Urban *et al.*, 1995).

Particularly, effects on Skills, Knowledge and Attitudes (SKAs) which fall under team cooperation and team coordination can be measured by capturing team communication. However, when communications among the team members are captured at a detailed semantic level, it is very difficult and time consuming to develop meaningful, quantitatively based measures to describe the nature of the communications. When, on the other hand, simple straightforward frequency counts of communications are used, no meaningful insight into team processes can be obtained.

The proposed approach, with an observation form (included in Appendix C) is between both aforementioned extremes and is a revision of the earlier Team Process Frequency (TPF) recording form. The revisions are twofold. First, the team behaviours categories (now types of verbal communications) to be captured have been kept to a minimum and are now hierarchically structured. Second, for each type of communication, the originator and the recipient are coded in accordance to Entin and Entin (2001).

Some of the rate measures that can be derived from the form are:

- total number of communications,
- number of requests for information,
- number of times information provided,
- number of requests for an action,
- number of announcements of actions,
- number of acknowledgements.

These rate measures can be captured both at the individual (team member) level and the team level. In order to compare data collected during trials or runs of different duration, these measures will be calculated per unit time, typically per minute, rather than as raw totals.

Anticipation ratios, measures based on the ratio of two of the aforementioned measures, have often proved more useful than single rate measures for understanding team communications (e.g. Entin & Serfaty, 1999).

The information anticipation ratio, the ratio of information announcements to information requests has proved particular useful for understanding team communications. Ratios larger than 1.0 are assumed to indicate that team members are anticipating the information needs and requirements of other team members and pushing them information before they request it. Ratios less than 1.0 are assumed to indicate that little anticipation of information needs is occurring and that team members must request (pull) the information they require from others.

The three useful and important anticipation ratios listed by Entin and Entin (2001) are:

- overall anticipation ratio – all announcements divided by all requests;
- information anticipation ratio – information announcements divided by information requests;
- action anticipation ratio - action announcements divided by action requests.

<u>Additional usability aspects of the observation form</u>

- When using video/audio recordings as backup, make sure that all communication between team members is recorded on video. This requires a camera position that captures all players of the team (preferably frontally, in order to see the face) and possibly separate microphones for each team member (but these can be recorded on the same audio channel).
- Radiotelephony (R/T) communication with aircraft is of no importance for the team communications recording form, but may be helpful for observers to understand team communication.
- When there are multiple observers using the recording form, they must use the same start time. This is not necessarily the start time of the scenario, but later, when the traffic picture has built up.
- The form is meant for recording task-related communication (i.e. communication about the tasks on hand). Non-task related communication (chatter) does not need to be recorded.
- In order not to interfere with team communication to be recorded on video and on the form, other persons present in the room need to be silent.

It needs to be checked:

- whether an observer can correctly complete the form without missing relevant communication;
- whether all relevant verbal communication within the team can be unambiguously be categorised within the available categories;
- the level of expertise required from the observer;

- the training that the observers need to have in order to reach an acceptable level of proficiency with the form (for a similar form, but within a military, possibly more complex organisation, Entin and Entin (2001) estimated 25 to 30 hours of training);

- whether a sufficient level of inter-observer reliability can be attained.

6. EXPERIMENTAL PLAN FOR VALIDATION OF THE TEAMWORK MEASURE

6.1 Introduction

SHAPE Teamwork Measure was performed in a host experiment that is part of a Commission of the European Communities (CEC) 5th framework project called 'Linking Existing ON Ground, ARrival and Departure Operations (LEONARDO)'[1]. In this section the experimental plan of that host will be briefly described. For more details about this experiment the reader is referred to the LEONARDO deliverables.

6.2 NLR Tower Research Simulator

NLR's real-time Tower Research Simulator (TRS) was used for the current study. The TRS is capable of simulating Tower Control activities at airports under nearly realistic operational conditions, with air traffic controllers and (pseudo) pilots in the control loop. The outside visual, provided with a projection screen of eleven by four meters, supports daylight and night-time situations, and winter conditions including snow and a variety of meteorological phenomena. The TRS has four working positions available for departure planner, ground controller, tower controller and gate operator. A photograph of the TRS is included in Figure 1.

[1] The LEONARDO Project (EU Contract number: GRD2/2000/30083) is funded by the Commission of the European Communities, Directorate General for Transport and Energy (DG-TREN) under the 'Competitive and Sustainable Growth' Programme (1998-2002).

The LEONARDO Consortium is coordinated by Entidad Pública Empresarial Aeropuertos Españoles y Navegación Aérea (EPE AENA, Spain) and further constitutes :

- Aéroports de Paris (ADP, France),
- Central Flow Management Unit (CFMU, EUROCONTROL),
- Iberia (Spain),
- INDRA Sistemas (Spain),
- Ingeniería y Economía del Transporte (INECO, Spain),
- Ingeniería de Sistemas para la Defensa de España (ISDEFE, Spain),
- The 'Ministère de l'Equipement, des Transports et du Logement', represented by the 'Direction de la Navigation Aérienne' (DNA, France),
- The Netherlands Aerospace Laboratory (NLR), and
- 'Sistemi Innovativi per il Controllo del Traffico Aereo' (SICTA, Italy).

Figure 1: NLR's Tower Research Simulator using the Amsterdam Schiphol airport database

6.2.1 The simulated airport

The simulated environment was the Amsterdam Schiphol airport. A detailed database of Schiphol and its environment was available. The simulated position of the tower was not exactly the same position as where the real tower is located. A location was selected that provided optimal view for the current experiment. Since the ATCOs were EUROCONTROL personnel, and not ATCOs who normally work in the Schiphol tower, this was not considered to have a negative impact on the current experiments.

6.2.2 The simulated aircraft

The tower research simulator was capable of simulating a total of eighteen different aircraft types and fifteen airport vehicles, both with various company paintings (i.e. liveries) are applied. Four pseudo pilots were simulating the flight crew for all simulated aircraft.

6.3 Air Traffic Controllers' Task

Three different air traffic controllers participated in the experiment. Each of them had a different task while together they formed a team that guided aircraft from gate to runway and the other way around from runway to gate. It was foreseen that aircraft waiting at the gate were handed off by the departure planner (also called 'start-up controller') first, and then via the ground controller were sent to the runway by the tower controller. Arriving aircraft were, via the tower controller, passed onto the ground controller and then at the gate were waiting to plan their next departure with the departure planner. So, in fact, the exchange of aircraft between the controllers was quite straightforward.

6.3.1 The automation

The main objective of the host experiment was to define a method and demonstrate the feasibility of integrating existing tools for arrival and departure planning, together with those derived from the planning and routing function of ground movement. This objective was achieved, amongst others, by performing a full-scale integration of existing ground, arrival and departure planning and management tools. The integration of these tools enabled the sharing of information between the tools, as well as their coordination through a functional ‘cooperative planning’ application layer. The results of the operational assessment provided a quantifiable measure of the safety, capacity and efficiency benefits of an integrated system. An example may be that the departure planner (the person, not the software tool) with CDM support switched on, received his schedules in a more optimal way because the CDM has amongst others taken the scheduled arrivals into account. This allowed the departure planner to focus his attention on the complex planning tasks that cannot be 'outsourced' easily to software.

The study will be complemented by a benchmarking study, in a simulated and generic environment. This simulated study is where SHAPE will validate its questionnaires.

The (NLR) tools that were integrated consisted of Arrival Manager (AMAN), Departure Manager (DMAN), Gate Allocation System (GAS), Taxi-in Planner (TIP) and Taxi-Out Planner (TOP). This integration is visualised in Figure 2.

The benchmarking study consisted of three conditions. The first was a reference system using close to present day procedures, with the tools operating independently, at the Schiphol airport tower. The second condition was called the Collaborative Decision-Making (CDM) system and comprises of the management tools (and actors) sending out their relevant information for use by other operators. The information flow of CDM is visualised in Figure 3. The third phase was called ‘Collaborative Decision-Making using Multiple Agent (CDMMA)’. This phase involved the tools taking decisions together by means of Multiple Agent Technology. This last condition will not be covered in the context of the SHAPE validation study. For SHAPE conditions 1 and 2 only will be compared.

As is visualised in Figure 2, the impact of the CDM automation definitely influences the departure planner, the person whose visual scanning behaviour will be measured.

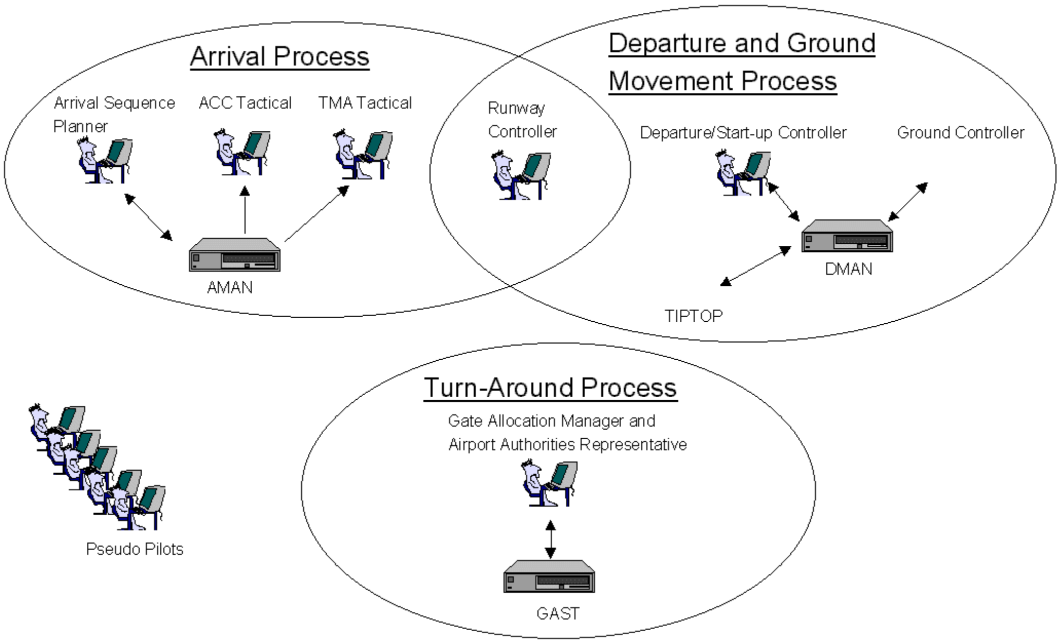


Figure 2: Graphical representation of condition 1 (reference condition)

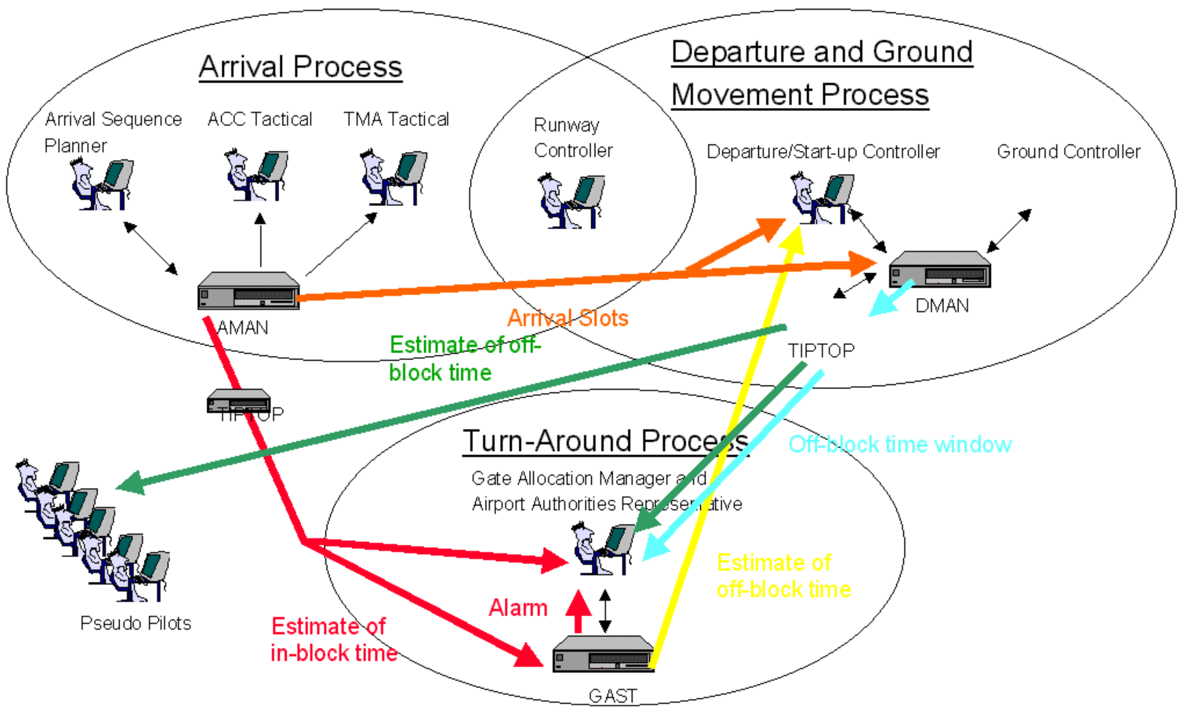


Figure 3: Graphical representation of condition 2 (with CDM information flow)

Figures 2 and 3 represent the information flow between the different actors in the host experiment. In addition, the physical distances and orientations of the air traffic controllers in the simulator are presented in Figure 4.

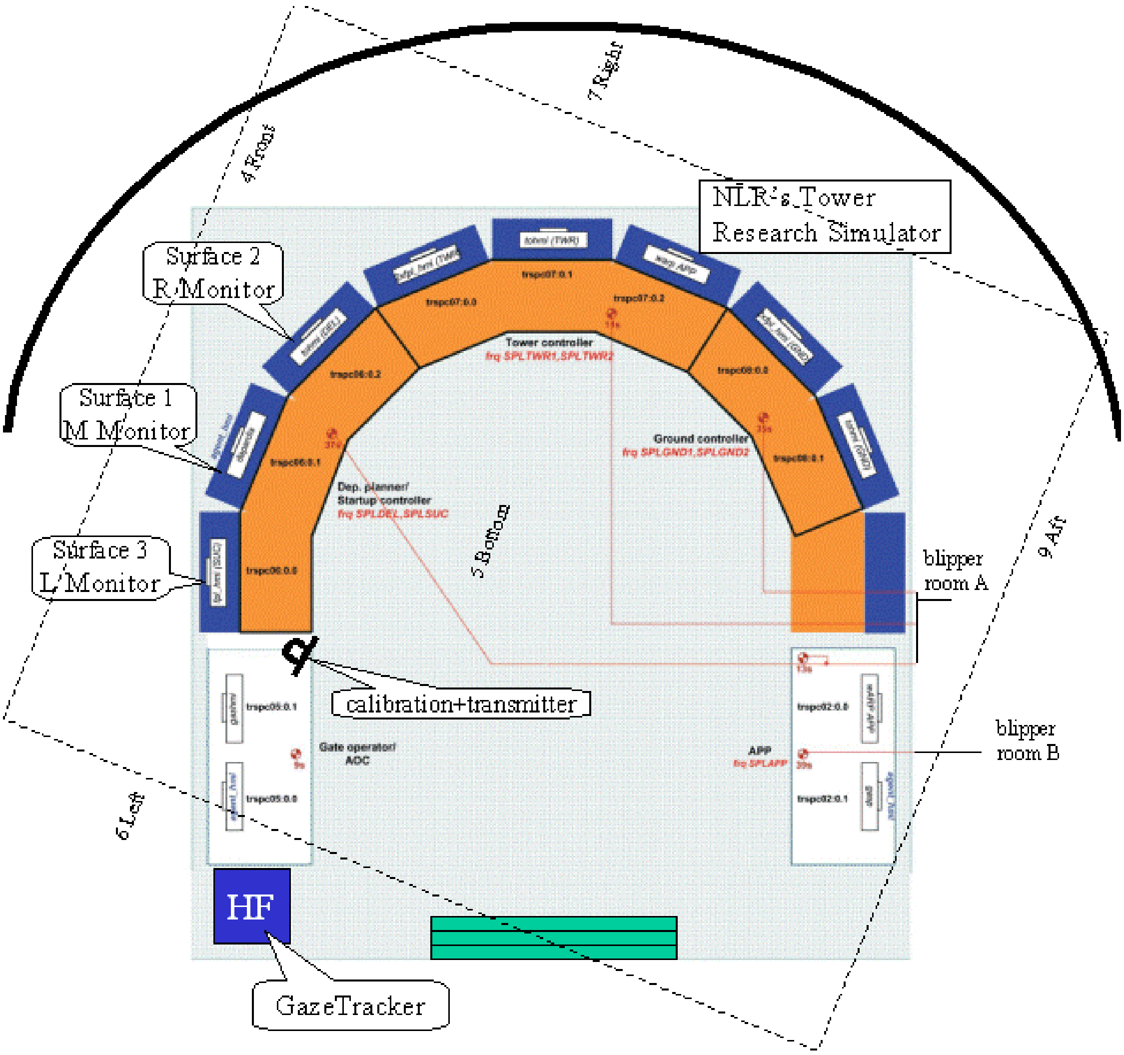


Figure 4: Physical distances and orientation of ATCOs working positions in the NLR TRS

Note: A representation of the surfaces monitored by the GazeTracker is also visualised in this figure. Some of the letters are difficult to read. The main aim of the figure is to give an overview of the positions in the Tower Research Simulator.

6.3.2 Communication task of air traffic controllers

Figure 5 shows an overview of the communication protocol for each ATCO in the experiment (with pilots, interface and other ATCOs). As can be seen the communication between ATCO and pilot is highly standardised.

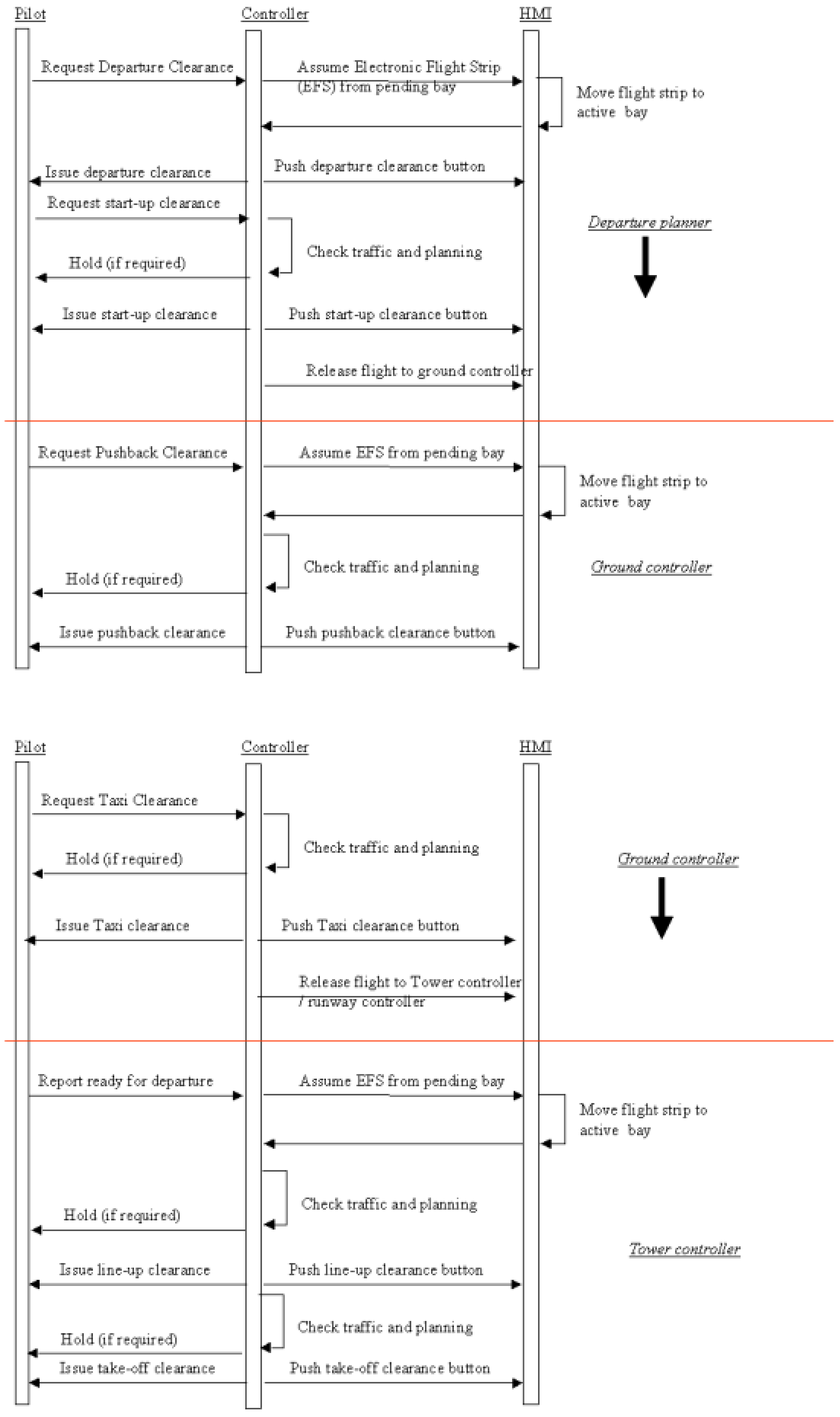


Figure 5: Communication protocol for each ATCO with pilots, interface and colleagues

6.3.3 The departure planners monitors

The departure planner - the ATCO whose eye scanning behaviour will be part of the current study - has three monitors in front of him. A brief description of the relevant aspects of function of each monitor for the departure planner in the current study will be given below. Note that the other ATCOs have monitors with comparable information and interface at their disposal.

Monitor 1 - Electronic Flight Strip (EFS) bays (see Figure 6). The EFSs in the middle of the monitor are under de departure planners control. The ones in the upper part of the monitor will become under his control when the pilot takes the initiative. The ones in the lower part of the monitor are finished and waiting to become under the control of the ground controller. On the right side of the monitor are visualisations of the required push back and taxi time in order to be at the runway on schedule.

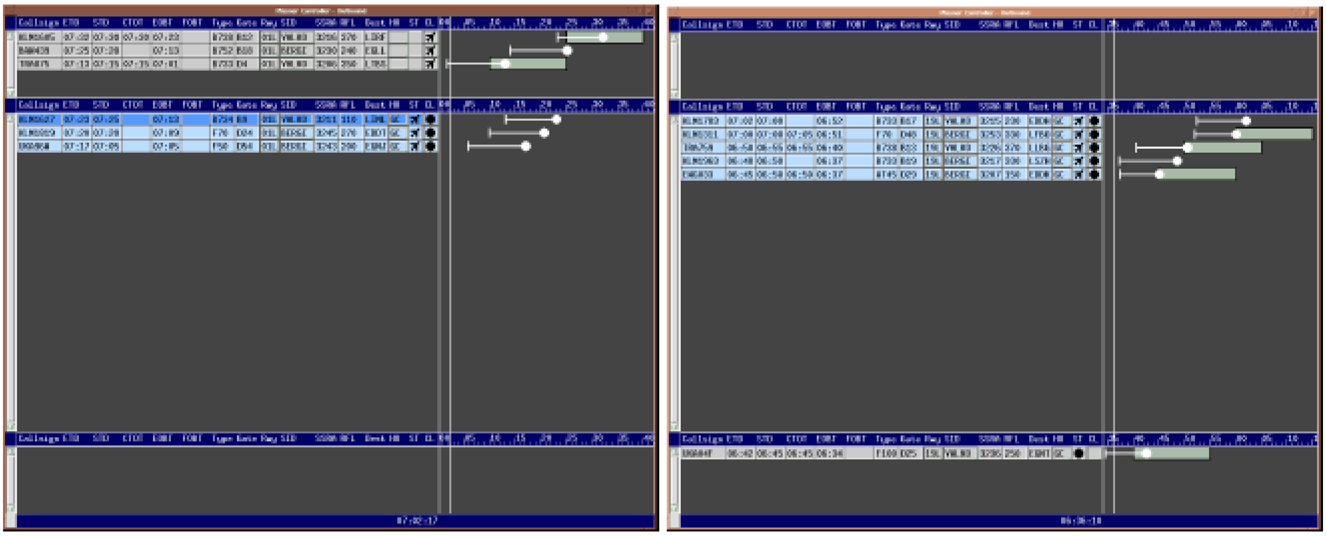


Figure 6: Examples of the monitors with EFSs (left - mixed runways, right - segregated runways)

Monitor 2 - Arrival and departure table(s) (see Figure 7). The yellow strips represent the arriving aircraft, the blue strips stand for the aircraft that are currently under de control of the departure planner and the grey strips are from aircraft that are not yet departed from the airport but who have already left the departure planners control.

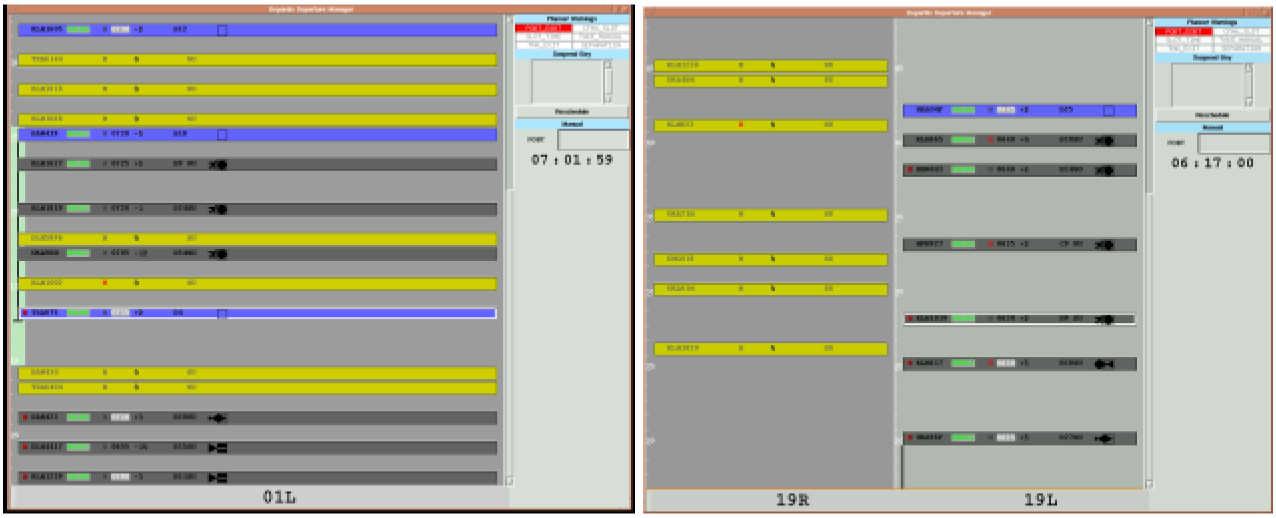


Figure 7: Examples of monitor with arrival and departure information (left - mixed runways, right - segregated runways)

Monitor 3 - Map of Schiphol airport's taxiways (see Figure 8). The yellow dots on the taxiways represent aircraft. As such they provide the departure planner with an impression of the business on the taxiways.

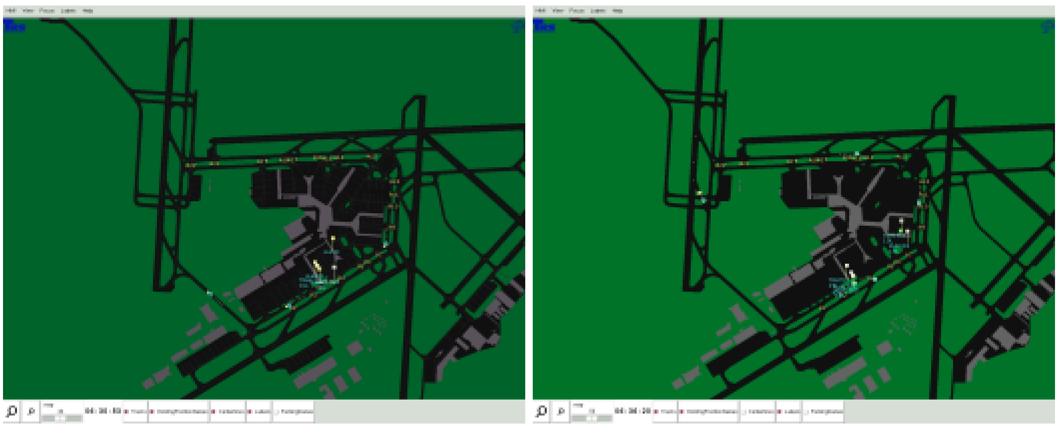


Figure 8: Map of traffic on taxiways around Schiphol airport (left - mixed runways, right - segregated runways)

Note: In the text and figures further in this document these monitors are also called:

- Monitor 1 = Electronic Flight Strip (EFS) bays = left monitor,
- Monitor 2 = arrival and departure table(s) = centre monitor,
- Monitor 3 = map of Schiphol airport's taxiways = right monitor.

6.4 Scenarios

There were four different scenarios. Each scenario was characterised by a number of departing and arriving aircraft. In half of the scenarios segregated runways for arrival and departure were used, while in the other half of the scenarios mixed runways were used. During the scenarios a number of key events took place. These were meant for the LEONARDO (host) experiment. For the current data analyses these were not used or disturbing the SHAPE data acquisition.

One of the differences between ‘segregated runway use’ and ‘mixed mode runway use’ that has impact on the current study is the fact that for segregated runways on the centre monitor of the departure planner two tables are presented. One table represents the arriving traffic while the other one represents the departing traffic. For the departure planner there is less (in some cases no) need, compared to the ‘mixed’ condition, to monitor the arriving traffic in order to line up his departing traffic in the most efficient way. However, in the mixed condition the moments when traffic is scheduled to arrive should be taken into account when making a schedule for departing traffic (from the same runway). When the CDM was operational it ‘helped’ the departure planner to schedule his departing traffic in an efficient way.

In all four scenarios only low to medium traffic density was applied. The reason for that is that the main goal of the LEONARDO studies was to examine the feasibility of the automation concept, so that a workload was required that would allow ATCOs to experiment with the tool and not to ‘overload’ them at this stage of concept validation already.

Since it was anticipated that the mixed mode runway use will be more demanding than the segregated runway use the scenarios for mixed run way use had a lower traffic density than the scenarios for segregated runway use. The expectation was that the mental workload for all four 'types of scenario' / 'use of runways' combinations would be about equal.

6.5 Experimental Design

The goal was to compare two levels of automation by confronting a team of air traffic controllers (i.e. a tower controller, a ground controller and a departure controller) in a simulated tower, with an automation-on and -off condition. The automation-on condition is also called the CDM-on condition. Further the experimental conditions will be varied by two different ways of runway use, and four different traffic scenarios. The experimental design is represented in Table 7.

Table 7:The experimental design in table format

	Scenario number	Condition 1 (reference)	Condition 2 (CDM)
Segregated runways	1S	run 1	run 5
	2S	run 6	run 4
Mixed runways	2M	run 3	run 2

Note: In the text and figures further in this document the conditions are labelled:

- R1REFS1 (run 1, reference/automation off, scenario: segregated runway use 1),
- R2CDMM2 (run 2, CDM/automation-on, scenario: mixed runway use 2),
- R3REFM2 (run 3, reference/automation-off, scenario: mixed runway use 2),
- R4CDMS2 (run 4, CDM/automation-on, scenario: segregated runway use 2),
- R5CDMS1 (run 5, CDM/automation-on, scenario: segregated runway use 1).

6.6 Measures and Equipment

A number of measures will be taken and a great deal of equipment will be installed in order to measure and record the data. Before going into detail about a number of these measures and the equipment that records the data, an overview will be presented. For a graphical representation see Figure 9.

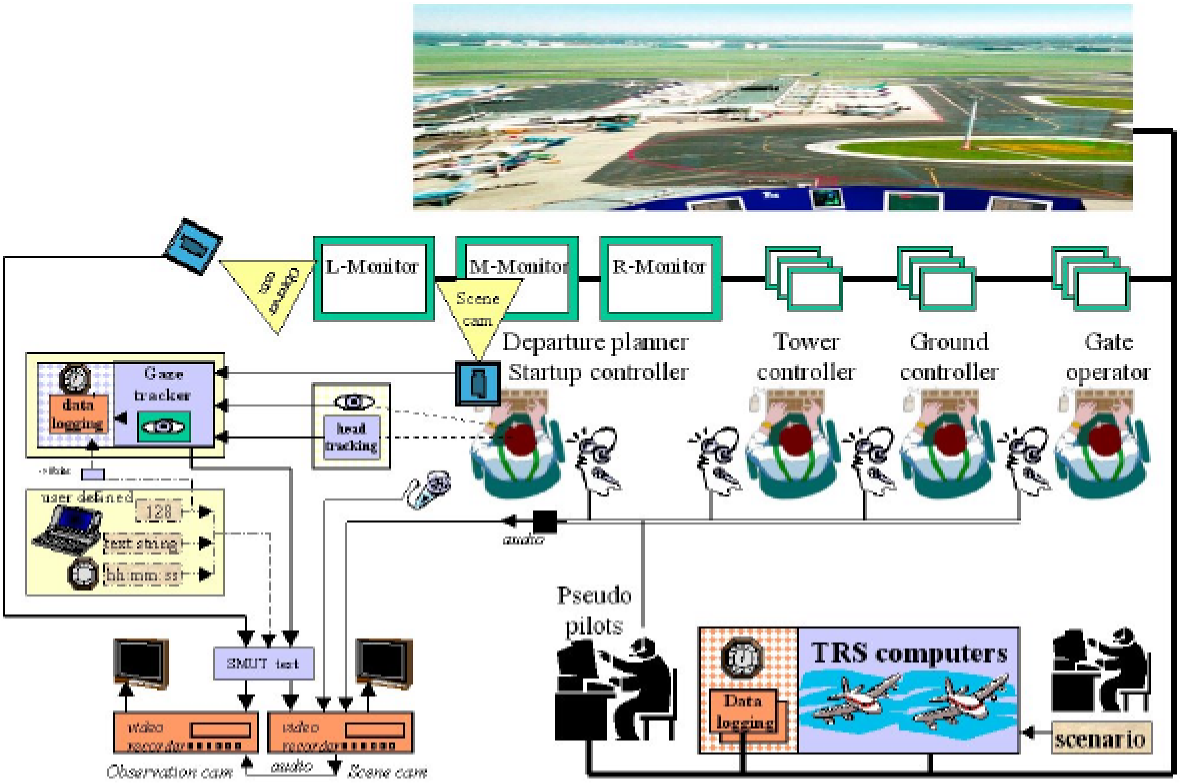


Figure 9: Graphical representation of the measurement equipment in NLR's Tower *research* Simulator

The scenario that runs on the simulator computers presents stimuli to the air traffic controllers. The pseudo pilots will act, within that scenario, as real pilots. The three air traffic controllers, and one member of the experiment team who will act as gate operator will be working in the simulated tower.

One fixed camera will record the entire team while another camera is mounted on the (eye tracker) helmet that will be worn by the departure controller. On this video the exact point of gaze of this departure controller will be visualised by a pair of superimposed cross hairs. Further the visual scanning behaviour of the controller will be recorded in a computer file. The questionnaires and rating scales that were developed for SHAPE will be used during, between and after the experimental runs.

6.6.1 Eye Point-of-Gaze

Eye Point-of-Gaze (EPOG) is the point on a predefined surface where an imaginary line coming straight from the centre of the eye crosses, via the pupil of the eye, that surface. As such this is the central point in the departure controllers field of vision. This point was measured by means of an EPOG recorder called GazeTracker (Mooij & Associates, 1996). An adjusted mountain-bike helmet was used to mount the optronics from the EPOG recorder that needs to be head mounted.

The duration that the departure controller gazes at a particular area of interest (see Figure 4) for the predefined Areas Of Interest [AOI]), is called the 'dwell time'. In addition to the dwell times, the scanning pattern, amounts of fixations, pupil diameter and eye-blink activity (that permits blink rate, duration and other measures to be derived) of the departure controller's left eyes were recorded as indicators of mental and visual workload (see Harris *et al.*, 1986; Stern *et al.*, 1984; Stern & Kelly, 1984; Stern, 1994; Goldstein *et al*., 1985; Wilson *et al.*, 1987, 1993). It is assumed that there is a negative correlation between (visual) workload and eye blink rate. The scanning behaviour is considered to be an indicator of the departure controller's mental state and focus of attention. It is generally assumed that if the departure controller looks at a particular AOI it means that he was paying attention to this area.

The different identified AIOs were:

- the three monitors:
  - *left monitor*: Electronic Flight Strips (EFSs),
  - *middle monitor*: arrivals and departures,
  - *right monitor*: overview of taxiways and runways on Schiphol airport;
- *outside*: the large wall representation of the outside view;
- the *desk*: the desk in front of the monitors where the controller read notes and other documentation; and finally
- the areas to the *left* and *right* of the departure planner as well as *behind* him, the *floor* and the *ceiling*.

The accuracy used to discriminate between fixations will be 1.5 degrees visual angle from the eye and a minimum fixation duration of 150 ms. For the analysis it will therefore be assumed that only if the departure controller looks at a point more than 1.5 degrees from the previous fixation, and for longer than 150 ms., a new fixation is made. For the determination of the eye blink rate, blinks that last longer then forty milliseconds will be considered as real blinks. Blinks are considered to be all further situations during which the pupil diameter is zero (closed eyelid).

Besides the digital data recording, EPOG was also recorded on video. The view filmed by a camera that was mounted on the departure controller's helmet and was aimed forward was mixed with a crosshair that was derived from the eye tracker. The crosshair indicated exactly where, in the camera image, the departure controller's eyes were aimed at.

For this analyses technique the definition of AOIs was slightly different from those used for the EPOG analysis. The definitions for the three *monitors* and for *outside* were the same, further *reading* (no matter where) was a different category, *looking over the shoulder* at colleagues was added and all additional scanning behaviour was rated as *other*.

Audio was recorded on the same tape. Everything that was said by the departure controller, the pilots that he spoke with and the other ATCOs were saying in the background, was recorded.

6.6.2 Rating scales and questionnaires

The previously described instruments that were developed as the SHAPE Teamwork Measure (Observation form (O), Questionnaire (Q), Self-rating form (R)) as well as the SHAPE Situation Awareness questionnaire (SASHA_Q) were used. For more information about SASHA the reader is referred to EATM (2003d).

6.6.3 Data collection

The data in the SHAPE/LEONARDO experiment were collected as follows. Three controllers participated in all runs. In each run they performed the same role. EPOG was recorded during each run, for the departure controller only. During each run the observation form was rated by external observers (i.e. members of the experiment team), after each run all three air traffic controllers completed the self-rating form, after all six runs the three air traffic controllers filled in the questionnaire.

6.6.4 Data analysis

Data analyses of questionnaires and EPOG data were performed using the Statistical Package for the Social Sciences (SPSS) release 10.0 and Microsoft Excel 97 sr 2. The EPOG data were first pre-processed with NLR's Gazeproc. software.

The videotapes were analysed in a software tool called ‘Ergoplayer’, an in house development of the Technical University of Munich (see Figure 10). For this analysis a so-called ‘task-oriented approach’ was chosen. This approach puts the emphasis of the analysis on the relationship between where the subject is looking and the particular part of the task that (s)he is working on at that moment. This results in specific information about eye scan patterns in the dynamic situations of working on (constantly changing) subsets of different tasks.

The pre-processing of the analysis comprised viewing the video tapes frame by frame and rating where the departure controllers point of regard was. This resulted in an Excel spreadsheet that gives information about where the departure controller’s point of regard was during each moment of the experiment. This analysis was done for runs 3 to 6.

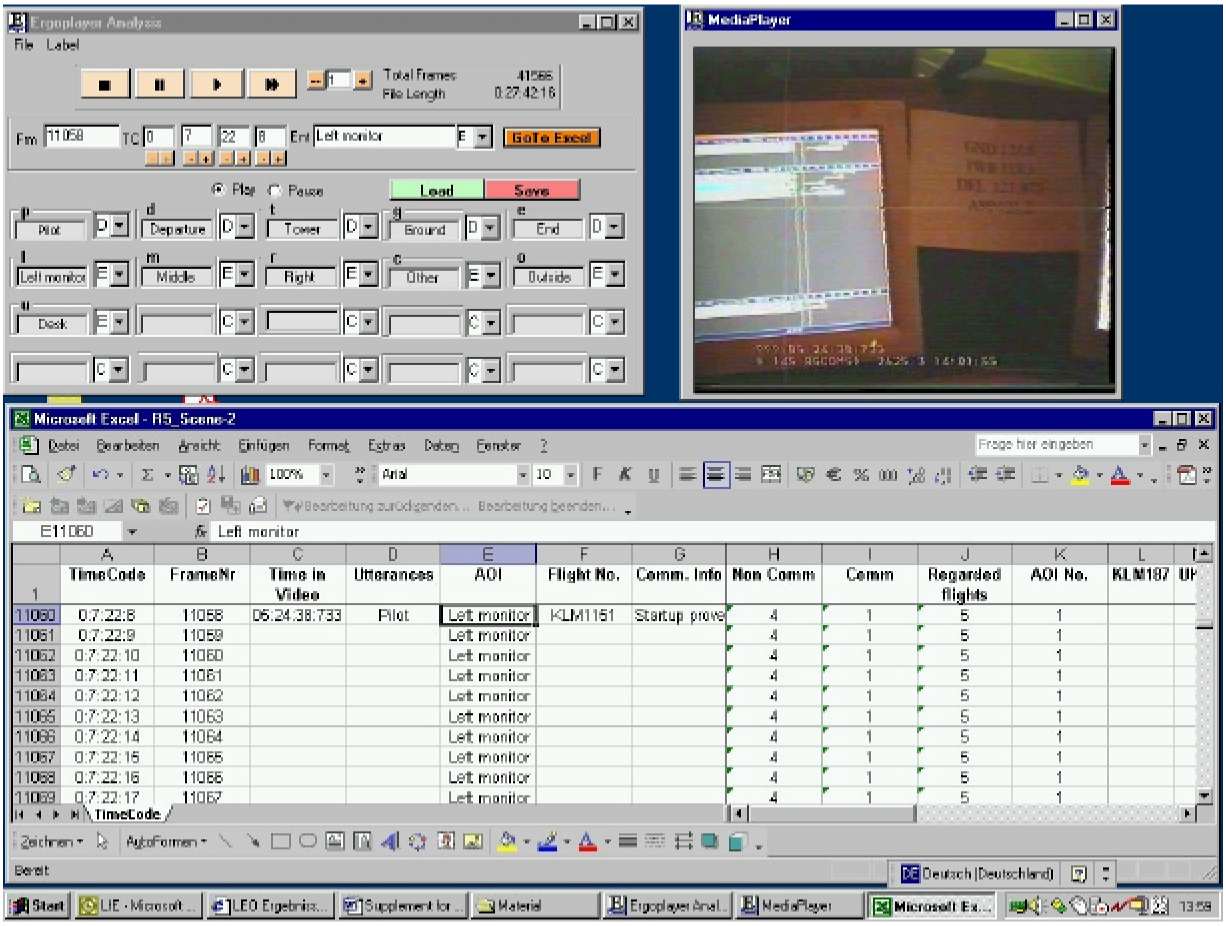


Figure 10: User interface of the ‘Ergoplayer’

Everything that was said by the departure controller and the pilots, who were under his control, was also logged, in a synchronised way, in the same spreadsheet. Finally some scenario related information that could be deduced from the context while viewing the videotapes was added. This comprised the amount of aircraft under control at each time. Based on these data, the 'task-oriented approach' analysis of Areas Of Interest (AOIs) was accomplished.

6.6.5 Validation procedure

In order to validate the questionnaire items, the procedure that is described below was used. The relationship between the SKATE Model constructs and questionnaire item, as was described in 5.5.4 of this document, was taken as a starting point. Hypotheses regarding the expected effect of the team performance in the measured variables were formulated for each item of the questionnaire. Separate hypotheses were formulated for those situations where automation was on or off. The next step was to compare the hypotheses with the ratings that the departure planner and ground controller had given on the questionnaire that was validated during the current experiment. The ratings that the controllers gave are visualised in Figure 11. Depending on the effect expected by the controllers, the researchers selected a hypothesis - measure combinations where this effect could be validated most effectively. This was done according to the 'task-oriented approach' (see 6.6.4). This expected effect on one or several measures formed the eventual hypotheses for each particular questionnaire item that will be validated in Section 7. The key measures used for verification/validation of the hypotheses were the communication (which was deduced from the videotapes afterwards), eye scanning behaviour and the self-rating form (R) (which consisted of previously validated scales). After the data analyses the hypotheses were either accepted or rejected and, as such items, validated or not on a three-point scale:

- Validated: The key measures fully support the rating of an item.
- Verified: The key measures do not contradict the rating of an item.
- Not validated: The key measures do contradict the rating of an item (because the analysis process turned out that the hypotheses would need to be more specific but constraints of the experimental design or data gathering process did not allow it).

For some cases it turned out to be impossible to validate them in the current experiment because of limitations concerning amounts of data, amounts of subjects or the experimental design. Because of these limitations it was decided to have, besides the categories 'validated' and 'not validated' a category 'verified'. 'Verified', in this context, means that it is likely that the hypothesis concerning this item is true, and that there are no reason to assume that the item cannot be validated, however, under the current conditions this was not possible. An example of such a situation is that no difference between automation on and off was expected because the ATCOs had rated that they expected no difference. If the data analysis proved that there was indeed no difference, this item was called 'verified'.

Note that the amounts of data will never allow to determine statistical significant differences but that only trends and tendencies can be used as indicators to validate or not.

A number of times there is no difference expected between condition A and B. Of course, according to the rules of statistics, ‘no difference’ may not be used as a condition to accept a hypothesis. For the current study, in mutual agreement with EUROCONTROL, it was decided that when the ATCOs had rated ‘neutral’, this would be interpreted as that no difference could be expected, and that this would be interpreted as inclining towards validation.

Page intentionally left blank

7. VALIDATION RESULTS

The current section describes the results of the validation study. Below are the 33 items of the team process questionnaire (Q) that was validated in this SHAPE teamwork questionnaire validation study. Items that were definitely validated are shown, as are those definitely not validated, or that could not be validated given limitations in experimental design or other analysis-related constraints. More details about the data upon which the validations are based may be found in Appendix D.

1. With the CDM system, it is clear to me which tasks are my own responsibility and have to be done manually.		VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 5	

2. With the CDM system, it is clear to me which tasks are done entirely by the system.		VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 2 Ground controller: 4	

3. With the CDM system, it is clear to me which tasks are done entirely by the other team members, with or without the system.		VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 5 Ground controller: 5	

4. With the CDM system, it is clear to me which tasks I share with the other team members.		VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 5 Ground controller: 4	

5. The CDM system allows the team to prioritise tasks more efficiently.		VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 2 Ground controller: 5	

6. With CDM, there is less need to ask other team members for assistance.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 5	

7. With CDM, there is less need to discuss working strategies with other team members.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 5	

8. My working style is compatible with the CDM system.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4

9. With the CDM system, my working style is no longer compatible with that of the other team members.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 3

10. CDM reduces the need for team communication.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 5

11. CDM promotes team spirit.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 4

12. With CDM, teamwork becomes more important.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 2

13. I have more trust in my team members when using the CDM system.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 3

14. With CDM, the goals of the team are more clearly defined.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 5

15. The team has the appropriate size for working with the CDM system.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 5 Ground controller: 4

16. With CDM, the responsibilities of the team members are clearly defined.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 5 Ground controller: 5

17. The team continues to learn to cooperate in working with the CDM system.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 4 Ground controller: 3	

18. The CDM system promotes clarity of roles in the team.		VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 2 Ground controller: 5	

19. The CDM system helps me to detect other team members' inaccuracies or mistakes.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 3	

20. The CDM system helps me to take initiative if required by the situation.		VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4	

21. The CDM system helps me to take command if required by the situation.		VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4	

22. The CDM system helps me to state my plans when required.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4	

23. The CDM system helps me to change plan, with other team members being informed/consulted.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 4	

24. With CDM, the task distribution among the team is appropriate.		VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 5 Ground controller: 5	

25. The CDM system helps me to intervene if task completion deviates from standards.		VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 4 Ground controller: 4	

26. The CDM system encourages the team to monitor each other.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 2 Ground controller: 2

27. The CDM system helps me sharing information about developing traffic situations with others.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 4

28. The CDM system helps team members alert each other to new events.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4

29. The CDM system helps the team to synchronise their actions so as not to hinder or delay each other.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 5

30. The CDM system makes it easier for the team to identify future overloading.	VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 2 Ground controller: 5

31. It is harder to motivate the team when working with the CDM system.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 3

32. When working with the CDM system, none of the team members needs to take command over the others.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4

33. With CDM, coordination between team members remains necessary.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 4

8. DISCUSSION AND CONCLUSIONS

The SHAPE Project's aim is to develop a measurement battery for human factors aspects of automation. In that context the current study has developed and validated a set of teamwork questionnaires. First a literature search was applied to gain up-to-date background knowledge about relevant teamwork aspects of automation. This information was fed into a theoretical model, which formed the base for the development of the set of questionnaires. Eventually the questionnaires were validated in a realistic simulator experiment.

Of course it is not just ATC where automation is an issue. Therefore the current study examined the field of automation in other domains as well. Special attention was given to:

- theories of teamwork in other domains, and
- impact of automation on teamwork and the assessment of this impact.

Besides differences between ATC and the other domains the following aspects of teamwork were identified as important to take into account for the current study:

- Attitudes of team workers towards teamwork.
- Communication and eye movement patterns should be examined in a dynamic way to find out more about (in)compatibility of information exchange. (This analysis technique is sometimes called a task-oriented approach.).
- Procedures and guidelines can be ordered according to 'severity' of the teamwork situation that needs to be automated.
- Dedicated training solutions for teamwork play an important role when introducing automation that has impact on teamwork.

Some of these aspects can be used as input for the development and validation of teamwork scales, some others can be helpful when interpreting the results that were measured with these scales.

Automation can take place at certain levels ranging from completely manual to completely automated. A one-dimensional scale for Level Of Automation (LOA) called 'Scale of Levels of Automation of Decision and Control Action' was taken as a starting point. Additional scales from NRC and EUROCONTROL were added to that scale. Eventually SHAPE identified a number of routine tasks for teams that could definitely be automated while a number of non-normal situations clearly cannot be automated.

SHAPE eventually distinguishes between six LOAs that are especially relevant for ATC:

1. Information extraction.
2. Information integration.
3. Information comprehension.
4. Decision/choice.
5. Response execution.
6. Information retention.

Based upon the information from ATC and other domains in combination with information about teamwork in general, and the SHAPE Project in particular, the SKATE Model was developed. This model is based upon four main dimensions:

1. Cooperation.
2. Coordination.
3. Attitudes towards teamwork.
4. Leader type qualities.

The SKATE Model items, which are representative for teamwork aspects of automation, are covered by means of three questionnaires that were developed in the current project. The questionnaires are:

- Observation form O: A form that allows (external) observers to rate a number of observable teamwork aspects that take place within the team when the team is performing their (automated) task.

- Questionnaire - Q: A form that the team members are supposed to fill in after being confronted with the new automation functionality. This form provides information of very specific, teamwork-related aspects of the impact of the particular type of automation that is being studied.

- Self-rating form – R: A form that is supposed to be filled in by team members and that provides information about workload and situation awareness aspects during a particular (simulator) work session. This is an ideal tool to verify the impact of the new automation functionality on workload and situation awareness in the context of the normal task.

The questionnaires were validated by applying them in a small-scale experiment that took place in NLR's Tower Research Simulator. Further data recording took place by filming the team, recording audio, and recording the eye scanning behaviour of one of the team members.

The self-rating form (R) consisted of previously validated scales. They were used as references against which questionnaire (Q) could be validated.

For each questionnaire (Q)-item hypotheses based upon the rating of the ATCOs on the questionnaire (Q) were formulated. These hypotheses regarded expected differences between the automation on and -off conditions with respect to the data recorded during the experiment.

The analysed eye tracker data provided interesting insights in the way that the departure planner divided his attention under the different conditions. This type of results has not very frequently let to the decision to validate/verify a questionnaire item. The reason for that is that this type of data often, and also in the current study, requires results from several subjects before solid statements can be made. An example of an interesting insight that was gained based upon eye tracker data is the fact that that departure planner made more fixations on the arrivals and departures tables under the mixed runway condition when the automation was switched off compared to when the automation was switched on. Possibly because the automation took care of a number of tasks that he otherwise had to deal with, and spent time on, himself (see Table 12 and Table 13 in Appendix D).

Eventually seventeen out of 33 items were either verified or validated. Therefore it can be concluded that at this moment the Questionnaire (Q) is partly validated. The first step in the process of validation was made. However, before this questionnaire can be applied as a stand-alone questionnaire in other experiments, a second step comprising more validation work and a reliability assessment is recommended.

For further validation of the SHAPE Teamwork Measure, it is suggested to develop an experiment that allows validation of the instruments that were not yet validated.

Teamwork in automated systems means more than communication. The study on a collaborative decision-making system in the tower simulator showed clearly that teamwork in automated environments means performance of distributed staff via technical systems. Normal communication channels, usually suitable for recovering from critical situations, are lost. As such Observation form (O) could not be validated because of the experimental setup of the current study. (The experimental setup interfered with natural communication between the ATCOs. This resulted in considerable less communication than in reality.)

Interviews with the ATCOs who participated in the experiment revealed that the workload of the communication partner could not be assessed in the way that they normally do. Consequently, the workload of the partner is miss-assessed and may lead to overload as well as critical situations. Similar statements could also be observed during datalink simulations, in particular in combination with system disturbances.

Automated systems cannot substitute speech act and need appropriate other means to overcome the loss of non-verbal communication (i.e. the richness of face-to-face communication in particular) regarding:

- delivery of emotional states,
- delivery of workload states,
- delivery of cognitive states (diagnosis of disturbances).

Designers of systems therefore have to carefully take into account the design and specification of new ATC systems regarding communication. This part of system design needs as much attention as the correct functioning of an isolated system. in fact one can say that, on top of the so-called user-centred design it becomes increasingly relevant and important to also take the team-centred design into account. The currently developed set of questionnaires is one mean that can be used to assure that system design of new automated ATC systems meet communication needs of controllers.

Finally, there are a number of constraints that need to be taken into account when interpreting the outcome of the current study.

1. With ever more automated systems there is less need for team members to exchange information in the classical way by voice and non-verbal communication. The less observable information exchange via monitors and input devices, that even may be quite indirectly via all kinds of software as intermediate filters, makes that new techniques to measure teamwork may be needed. It also makes that we must prohibit to be mislead by just clearly observational exchange of information. As such information exchange is ever more becoming a system task, rather than solely the domain of team members. This development should definitely be part of future teamwork questionnaires. As a result the amount of data needed to validate the Observation form (O) was not sufficient and no effort was put in attempts to validate that form in the current study.

2. The design of the current experiment was developed for other purposes than a validation study. In this design it was decided that the ATCOs who normally exchange most of the information would sit furthest away from each other. This made it more difficult for the ATCOs to perform a normal information exchange. The ATCOs have also confirmed this repeatedly after the experiment. This way of operation is, of course, contrary to a situation that is as normal as possible with only one variable (automation) being manipulated, as would be desirable for the current validation study. Therefore there may have been less information exchange in the current study than would have been observable under more realistic conditions.

3. Only part of the teamwork measures could be validated, due to a limited number of participants, not always having suitable validating measures and budgetary restrictions. Therefore, validation in the statistical sense, e.g. rejection of general hypotheses about the impact of automation on teamwork with a certain percentage of confidence, was impossible. All results are based upon trends.

REFERENCES

Avermaete, van J. & Kruijsen, E. (Eds) (1998). The Evaluation of Non-Technical Skills of Multi-Pilot Aircrew in Relation to the JAR-FCL Requirements. *NLR Technical Publication NLR TP-98-518*, National Aerospace Laboratory, Amsterdam, The Netherlands: NLR.

Baker, D.P. & Salas, E. (1992). Principles for measuring teamwork skills, *Human Factors*; 34-4, 469-475.

Bohnen, H.G.M. (2003). Calculation of team workload for the NH90 - Evaluation of methodology. *Internal NLR communication*.

Bowers, C.A., Deaton, J., Oser, R., Prince, C. & Kolb, M. (1993). The impact of automation on crew communication and performance. In: R. Jensen (Eds), *Proceedings of the Seventh International Symposium on Aviation Psychology*, 758-761, Columbus, OH, US: Ohio State University.

Bowers, C.A., Oser, R.L, Salas, E. & Cannon-Bowers, J.A. (1996). Team performance in automated systems. In: R. Parasuraman & M. Mouloua (Eds), *Automation and Human Performance: Theory and Applications*, 243-263, Mahwah, US: Lawrence Erlbaum.

Cannon-Bowers, J.A., Tannenbaum, S.I., Salas, E. & Volpe, C.E. (1995). Defining team competencies and establishing team training requirements. In: R. Guzzo & E. Salas (Eds), *Team effectiveness and decision-making in organizations*, 333-380. San Francisco, CA, US: Jossey Bass.

CAST (1998a). Consequences of future ATM systems for air traffic controller Selection and Training (CAST). WP1: Current and Future ATM systems. *NLR Technical Report NLR-TR-98233*. Amsterdam, The Netherlands: NLR.

CAST (1998b). Consequences of future ATM systems for air traffic controller Selection and Training (CAST). WP2: Future ATCO description. *NLR Technical Report NLR-TR-98496*. Amsterdam, The Netherlands: NLR.

Chung, G.K.W.K., O'Neil, H.F. & Herl, H.E. (1999). The use of computer-based collaborative knowledge mapping to measure team processes and team outcomes. *Computers in Human Behavior*, 15-3-4, 463-493.

Crichton, M. & Flin, R. (1999). Training of non-technical skills for emergency management – UK nuclear power industry experience. *Proceedings of the 2nd International Disaster and Emergency Readiness Conference*, October. The Hague, NL: Andrich International.

Cox, M. (1994). Task analysis of selected operating positions within UK air traffic control. Vol. 1, Main Report, and Vol. 2, Appendices. *Royal Air Force Institute of Aviation Medicine Report No. 749*, Farnborough, Hants.

Dickinson, T.L. & McIntyre, R.M. (1997). A conceptual framework of teamwork assessment. In: M.T. Brannick, E. Salas & C. Prince (Eds), *Team performance assessment and measurement: Theory, methods and applications*, 331-356. Mahwah, NJ: Erlbaum.

Driskell, J.E. & Salas, E. (1992). Collective behavior and team performance. *Human Factors*, 34, 277-288.

Dwyer, D.J. (1984). Team research and team training: a state of the art review. In: F.A. Muckler (Ed.), *Human Factors Review.* Santa Monica, CA, US: Human Factors and Ergonomics Society.

Dwyer, D.J., Fowlkes, J.E., Oser, R.L., Salas, E. & Lane, N.E. (1997). Team performance measurement in distributed environments: the TARGETs methodology. In: M.T. Brannick, E. Salas & C. Prince (Eds), *Assessment and Measurement of Team Performance: Theory, Research and Application*, 137-154. Hillsdale, US: Lawrence Erlbaum.

EATCHIP Human Resources Team (1996a). *Awareness of Human Aspects in Air Traffic Services – Guidance for Managing Current Issues*. HUM.ET1.ST13.1000-REP-01. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATCHIP Human Resources Team (1996b). *Human Factors Module - Stress*. HUM.ET1.ST13.2000-REP-01. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATCHIP Human Resources Team (1997a). *Human Factors Module – Interpersonal Communication*. HUM.ET1.ST13.1000-REP-02. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATCHIP Human Resources Team (1997b). *Human Factors Module – Critical Incident Stress Management*. HUM.ET1.ST13.3000-REP-01. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATCHIP Human Resources Team (1998a). *Integrated Task and Job Analysis of Air Traffic Controllers – Phase 1: Development of Methods*. HUM.ET1.ST01.1000-REP-03. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATCHIP Human Resources Team (1998b). *Human Factors Module – Human Factors in the Investigation of Accidents and Incidents*. HUM.ET1.ST13.3000-REP-02. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATCHIP Human Resources Team (1998c). *Human Factors Module – Human Factors in the Development of Air Traffic Management Systems*. HUM.ET1.ST13.4000-REP-01. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATCHIP Human Resources Team (1998d). *Proceedings of the Second EUROCONTROL Human Factors Workshop - Teamwork in Air Traffic Services*. HUM.ET1.ST13.000-REP-02. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATMP Human Resources Team (1999a). *Integrated Task and Job Analysis of Air Traffic Controllers – Phase 2: Task Analysis of En-route Controllers.* HUM.ET1.ST01.1000-REP-04. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATMP Human Resources Team (1999b). *Human Factors Module - A Business Case for Human Factors Investment*. HUM.ET1.ST13.4000-REP-02. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATMP Human Resources Team (1999c). *Human Factors Module – Learning and Skills Acquisition*. HUM.ET1.ST07.1000-REP-01. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATMP Human Resources Team (1999d). *Team Resource Management Test and Evaluation*. HUM.ET1.ST10.2000-REP-01. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATMP (2000). *Human Resources Programme – Stage 1: Programme Management Plan*. Edition 1.0. Brussels: EUROCONTROL.

EATM Human Resources Team (2003a). *Guidelines for Trust in Future ATM Systems: A Literature Review*. HRS/HSP-005-GUI-01. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATM Human Resources Team (2003b). *Guidelines for Trust in Future ATM Systems: Measures*. HRS/HSP-005-GUI-02. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATM Human Resources Team (2003c). *Guidelines for Trust in Future ATM Systems: Principles*. HRS/HSP-005-GUI-03. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATM Human Resources Team (2003d). *The Development of Situation Awareness Measures in ATM Systems*. HRS/HSP-005-REP-01. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATM Human Resources Team (2003e). *Age, Experience and Automation in European Air Traffic Control*. HRS/HSP-005-REP-02. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATM Human Resources Team (2004a). *Impact of Automation on Future Controller Skill Requirements and a Framework for their Prediction*. HRS/HSP-005-REP-04. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATM Human Resources Team (2004b). *Managing System Disturbances in ATM: Background and Contextual Framework*. HRS/HSP-005-REP-06. Edition 1.0.. Released Issue. Brussels: EUROCONTROL.

EATM Human Resources Team (2004c). *A Tool for the Assessment of the Impact of Change in Automated ATM Systems on Mental Workload*. HRS/HSP-005-REP-03. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

EATM Human Resources Team (2004d). *Age, Experience and Automation in European Air Traffic Control*: *Survey in the ECAC Area*. HRS/HSP-005-REP-05. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

Endsley, M.R. (1997). Level of automation: integrating humans and automated systems. In: *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting.* Santa Monica, CA, US: Human Factors Society.

Eissfeldt, H. (1988). Beurteilungsrelevante Leistungskriterien im Rahmen der praktischen Ausbildung zum gehobenen Flugverkehrskontrolldienst - Eine Untersuchung anhand der Critical Incident Technique. DLR-IB-316-88-07.

Eissfeldt, H. (1991). Automation in ATC: How does it affect the Selection of Controllers? In: J.A. Wise, V.D. Hopkin & M.L. Smith (Eds), *Automation and System Issues in Air Traffic Control*, Vol. NATO ASI Series F, 73, 461-465. Berlin: Springer.

Eissfeldt, H. (2002). Using a computer-based work sample following the Lerntest-approach to predict the success in practical phases of air traffic controller training (*in German*). *DLR-Forschungsbericht 2002-13*. Germany: DLR.

Entin, E.E. (1999). Optimized command and control architectures for improved process and performance. *Proceedings of the 1999 Command and Control Research and Technology Symposium*, Newport, RI.

Entin, E.E. & Entin, E.B. (2001). Measures for evaluation of team processes and performance in experiments and exercises. In: *Proceedings of the 6th International Command and Control Research Technology Symposium*, June 19-21, Annapolis, MD, US.

Entin, E.E. & Serfaty, D. (1999). Adaptive team co-ordination. *Human Factors*, 41-2, 312-325.

ESSAI (2000). Enhanced Safety through Situation Awareness Integration in training. WP1: Orientation. *NLR Technical Report NLR-TR-2000-668*. NLR: Amsterdam, The Netherlands.

ESSAI (2003). Enhanced Safety through Situation Awareness Integration in training. WP6: Final Report. *NLR Contract Report NLR-CR-2003-064*. NLR: Amsterdam, The Netherlands.

EUROCONTROL PHARE PD/1 (1997). Programme for Harmonised Air Traffic Management Research in EUROCONTROL, PD/1, Final Report, PHARE/NATS/PD1-10.2/SSR;1.1 - *EUROCONTROL DOC 96-70-24*. Available from http://www.eurocontrol.int.

EUROCONTROL PHARE PD/3 (1998). Programme for Harmonised Air Traffic Management Research in EUROCONTROL, PD/3, Ground Human-Machine Interface Specification [V2.2], PHARE/EEC/GHMI-6.7.1. Brussels: EUROCONTROL.

FAA (1996). *Data Link Benefits Study Team Report*. Washington DC, US: US Department of Transportation.

Fitts, P. (1951). Engineering Psychology and Equipment Design. In: S. Stevens (Ed.), *Handbook of Experimental Psychology.* 1287-1340. New York, US: John Wiley & Sons, Inc.

Fletcher, G., McGeorge, P., Flin, R., Glaving, R. & Maran, N. (2002). The role of non-technical skills in anaesthesia: A review of current literature. *British Journal of Anaesthesia*, 88, 418-429.

Flin (1996). *Sitting in the Hot Seat: Leaders and Teams for Critical Incident Management*, Chichester, UK: John Wiley.

Flin, R., Fletcher, G., McGeorge, P., Sutherland, A. & Patey, R. (2003). Anaesthetists' attitudes to teamwork and safety. *British Journal of Anaesthesia*, 58(3):233-42.

Flin, R. & Slaven G. (1994). *The selection and training of offshore installation managers for Crisis Management*. London, UK: HSE-Books.

Flin, R., Slaven, G. & Stewart, K. (1996). Emergency Decision-making in the Offshore Oil and Gas Industry. *Human Factors*, 38(2), 262-277.

Fukuda, R., Voggenberger, T., Sträter, O. & Bubb, H. (2003). *Measuring the quality of group interaction in nuclear power plant environments by using eye movement behavior*. Berlin, Germany: GIHRE.

Gaba, D.M., Fish, K.J. & Howard, S.K. (1994). *Crisis Management in Anaesthesiology*. New York, US: Churchill Livingstone.

Goldstein, R., Walrath, L.C., Stern, J.S. & Strock, B.D. (1985). Blink activity in a discrimination task as a function of stimulus modality and schedule of presentation. *Psychophysiology*, 22, 629 - 635.

Gregorich, S.E., Helmreich, R.L. & Wilhelm, J.A. (1990). The structure of cockpit management attitudes, *Journal of Applied Psychology*, 75, 782-690.

Hackman, J.R. (1990). *Groups that work (and those that don't): Creating conditions for effective teamwork*. San Francisco, US: Jossey-Bass.

Hanson, E., Zon, R. & Hilburn, B. (2003). Losing Situation Awareness The first symptoms. In: R. Jensen (Ed.) *Proceedings of the International Symposium of Aviation Psychology*, Dayton, OH, US.

Harris, R.L., Glover, B.J. & Spady, A.A. (1986). Analytical Techniques of Pilot Scanning Behavior and Their Application*. NASA Hampton Technical Paper 2525*.

Hart, S.G. & Staveland, L.E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In: P.A. Hancock & N. Meshkati (Eds), *Human mental workload*, 139-183. Elsevier Amsterdam, The Netherlands.

Helmreich, R.L., Foushee, H.C., Benson, R. & Russini, R. (1986). Cockpit Resource Management: Exploring the attitude-performance linkage. *Aviation, Space and Environmental Medicine*, 57, 1198-1200.

Helmreich, R.L., Merritt, A.C., Sherman, P.J., Gregorich, S.E. & Wiener, E.L. (1993). The Flight Management Attitudes Questionnaire (FMAQ). *NASA/UT/FAA Technical Report 93-4*. Austin, TX, US: The University of Texas.

JAR-TEL (2001). Consolidation of Results, *WP7 Draft Report, JAR-TEL /NLR/WP7/D9_02*, EU Sponsored Framework V Project, Contract n° AI-97-SC.2037.

Jentsch, F., Bowers, C., Redshaw, B., Bergen, H., Henning, J. & Holmes, B. (1995). Differential effects of automation on team performance, workload and communications. Technical Report, Naval Air Warfare Centre - Training Systems Division, FL, US.

Jorna, P. (1997). Human-machine interactions with future flight deck and air traffic control systems. In: D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics*. UK: Ashgate Publishing.

Kerns, K. (1994). Human Factors in ATC / Flight Deck Integration: Implications of Data Link Simulation Research. Report Mitre Corporation. VA, US: McLean.

Klein, G. (1989). *Recognition-primed decisions. Advances in man-machine systems research*, 5, 47-92.

Kleinman, D.L. & Serfaty, D. (1989). Team Performance assessment in distributed decision-making. *Proceedings of the Symposium on*

Interactive Networked Simulation for Training, 22-27, Orlando, FL, US.

Klimoski, R. & Mohammed, S. (1994). Team Mental Model: Construct or Metaphor? *Journal of Management*, 20-2, Summer, 403-437.

Langan-Fox, J., Wirth, A., Code, S., Langfield-Smith, K. & Wirth, A. (2001). Analyzing shared and team mental models. *International Journal of Industrial Ergonomics*, 28-2, 99-112.

Masson, M. & Paries, J. (1998). Team Resource Management Training for Air Traffic Controllers, in: *Proceedings of the Second EUROCONTROL Human Factors Workshop. - Teamwork in Air Traffic Services*, EUROCONTROL, Luxembourg. HUM.ET1.ST13.000-REP-02. Edition 1.0. Released Issue. Brussels: EUROCONTROL. (available from http://www.eurocontrol.int/humanfactors/publ.html).

McIntyre, R.M. & Salas, E. (1995). Measuring and managing for team performance: Emerging principles from complex environments. In: R. Guzzo & E. Salas (Eds), *Team effectiveness and decision-making in organizations*, 149-203. San Francisco, CA, US: Jossey-Bass.

Meckiff, C. & Gibbs P. (1994). PHARE Highly Interactive Problem Solver. In: *EUROCONTROL Experimental Centre (EEC) Report 273/94*. Brétigny, France: EEC.

Meister, D. (1985). *Behavorial measurement and analysis methods*. New York, US: Wiley.

Modrick, J.A. (1986). Team performance and training. In: J. Zeidner (Ed.), *Human Productivity Enhancement: Training and Human Factors in Systems Design*, Vol. 1. New York, US: Praeger.

Mooij & Associates (Zon, G.D.R. & Mooij, H.A.) (1996). GazeTracker for Real-time, Omni-directional and Digital Point-of-gaze Determination - Installation and Operation Guide - EPOG Version 4.3, MA-95-15, December.

Morgan, B.B. Jr., Glickman, A.S., Woodward, E.A., Blaiwes, A.S. & Salas, E. (1986). Measurement of team behaviours in a navy environment. *NTSC technical report TR-86-014*, Orlando, FL, US: NTSC.

Mosier, K.L. & Skitka, L.J. (1996). Human decision-making and automated decision aids: made for each other? In: R. Parasuraman & M. Mouloua (Eds), *Automation and Human Performance: Theory and Applications*, 201-220. Mahwah, US: Lawrence Erlbaum.

Motamedzade, M., Shahnavaz, S., Kazemnejad A, Azar, A. & Karimi, H. (2002). *Assessment of team working strategy for implementing ergonomics at two production industries in Iran*. Paper presented at the second international cyberspace conference on Ergonomics. Paper available from: http://cyberg.wits.ac.za/.

Mullen, B. & Copper, C. (1994). The relation between group cohesiveness and performance: an integration. *Psychological Bulletin*, 115, 210-217.

Nieva, V.F., Fleishman, E.A. & Reick, A. (1978). *Team Dimensions: Their Identity, Their Measurement, and Their Relationships*. D.C.: Response Analysis Corporation.

NOTECHS (1998). Final report: The Evaluation of non-technical skills of multi-pilot aircrew in relation to the JAR-FCL requirements. *NLR Contract Report NLR CR-98443*, National Aerospace Laboratory, NLR: Amsterdam, The Netherlands.

NRC (1998). *The future of air traffic control - Human operators and automation*, by C.D. Wickens, A.S. Mavor, R. Parasuraman & J.P. Magee (Eds). National Research Council (NRC), Washington DC, US: National Academy Press.

Orasanu, J.M. (1990). Shared mental models and crew decision-making, *CSL report 46*, Cognitive Science Laboratory, NJ, US: Princeton University.

Orasanu, J. & Fischer U. (1997). Finding decisions in natural environments, the view from the cockpit. In: C. Zsambok & G. Klein (Eds), *Naturalistic decision making*. Hillsdale, NJ, US: LEA.

Parasuraman, R., Sheridan, T.B. & Wickens, C.D. (2000). A model for types and levels of human interaction with automation. IEEE Transactions on Systems, *Man and Cybernetics*. Part A: Systems and Humans, 30(3), 286-297.

Paris, C.R., Salas, E. & Cannon-Bowers, J.A. (2000). Teamwork in multi-person systems: a review and analysis. *Ergonomics*, 43-8, 1052-1075.

Pearn, M.A. & Kandola, R.S. (1983). *Job Analysis of Air-Traffic Controllers*. Esher, Surrey: SHL.

Porterfield, D.H. (1997). Evaluating controller communication time as a measure of workload. *The International Journal of Aviation Psychology*, 7(2), 171-182.

Prince, C. & Salas, E. (1993). Training and research for teamwork in the military aircrew. In: E.L. Wiener, B.G. Kanki & R.L. Helmreich (Eds), *Cockpit resource management*, 337-366. Orlando, FL, US: Academic Press.

Ruffell-Smith, H.P. (1979). A simulator study of the interaction of pilot workload with errors. *NASA Technical Report No. TM-78482*. Moffett

Field, CA, US: National Aeronautics and Space Administration-Ames Research Center.

Salas, E. & Cannon-Bowers, J.A. (2000). The anatomy of team training. In: L. Tobias & D. Fletcher (Eds), *Handbook on research in training*. New York, US: MacMillan.

Salas, E., Dickinson, T.L., Converse, S. & Tannenbaum, S.I. (1992). Towards an understanding of team performance and training. In: R. Swezey & E. Salas (Eds), *Teams, their training and performance*, 3-29. Ablex, Norwood.

Sheridan, T.B. & Verplank, W.L. (1978). Human and computer control of undersea teleoperators. *Technical report*. MIT Man-machine systems laboratory, Cambridge, MA, US.

Sherman, P.J. & Helmreich, R.L. (1995). Attitudes toward automation: The effect of national culture. In: *Proceedings of the Eighth International Symposium on Aviation Psychology,* 682-687. Columbus, OH: Ohio State University.

Smith, K. (2000) Adult Literacy and Lifeskills Survey (ALL): Teamwork Content. Statistics Canada.

Smith-Jentsch, K.A., Zeisig, R.L., Acton, B. & McPherson, J.A. (1998). Team dimensional training: A strategy for guided team self-correction. In: J.A. Cannon-Bowers & E. Salas (Eds), *Making decisions under stress: Implications for individual and team training*, 271-297 Washington, DC, US: APA Press.

Smith-Jentsch, K.A., Salas, E. & Baker, D. (1996). Training team performance-related assertiveness. *Personnel Psychology*, 49, 909-936.

Stern, J.A. (1994). Blink Rate: A Possible Measure of Fatigue. *Human factors*. 36(2), 285-297.

Stern, J.A. & Skelly, J.J. (1984). The eyeblink and workload considerations. *Proceedings of the Human Factors Society 28th Annual Meeting*, San Antonio, TX, US.

Stern, J.A., Walrath, L.C. & Goldstein, R. (1984). The Endogenous Eyeblink. *Psychophysiology*. 21(1), 22-33.

Urban, J.M., Bowers, C.A., Cannon-Bowers, J.A. & Salas, E. (1995). The importance of team architecture in understanding team processes. In: M. Beyerlein, D.A. Johnson & S.T. Beyerlein (Eds), *Advances in interdisciplinary studies of work teams*, 2, 205-228. Greenwich, CT: JAI Press.

VINTHEC I (1999). Final Report: Visual INTeraction and Human Effectiveness in the Cockpit I, *VINTHEC-WP9-TR 01*, Defence Evaluation and Research Agency, DERA Farnborough Hants, UK.

VINTHEC I (2000). Full-scale studies: Visual INTeraction and Human Effectiveness in the Cockpit I. EU-Sponsored Project. *NLR Technical Report NLR-TR-99404*, National Aerospace Laboratory, NLR, Amsterdam, The Netherlands.

VINTHEC II Consortium (2003). Final Report: Visual INTeraction and Human Effectiveness in the Cockpit II. EU-Sponsored Project. Report Available from http://www.nlr.nl/public/hosted-sites/vinthecII/index.html.

Volpe, C.E., Cannon-Bowers, J.A., Salas, E. & Spector, P. (1996). The impact of cross training on team functioning. *Human Factors*, 38, 87-100.

Wickens, C.D. (1994). Designing for situation awareness and trust in automation. In: *Proceedings of the IFAC conference on Integrated Systems Engineering*. Laxenburg, Austria: International Federation of Automatic Control.

Wiener, E.L. (1993). Crew Co-ordination and training in the advanced technology cockpit. In: E.L. Wiener, B.G. Kanki & R.L. Helmreich (Eds), *Cockpit Resource Management*, 199-229. San Diego, CA, US: Academic Press.

Wilson, G.F., Badeau, A. & Gawron, V. (1993). Evaluation of two inflight sensors using heart-rate and eye-blink measures. *Aviation Psychology*.

Wilson, G.F., Purfis, B., Skelly, J., Fullenkamp, P. & Davis, I. (1987). Physiological data used to measure pilot workload in actual flight and simulator conditions. In: *Proceedings of Human Factors society - 31st annual meeting*.

Zijlstra, F.R.H. (1993). *Efficiency in work behavior. A design approach for modern tools*. PhD thesis, Delft University of Technology. Delft, The Netherlands: Delft University Press.

FURTHER READING

EATMP Human Resources Team (2002). *Core Requirements for ATM Working Positions: An Overview of the Project Activity*. HRS/HSP-006-REP-03. Edition 1.0. Released Issue. Brussels: EUROCONTROL.

ABBREVIATIONS AND ACRONYMS

For the purposes of this document the following abbreviations and acronyms shall apply:

ACC	Area Control
ACP	Access Control Point
ACRM	Anaesthesia Crisis Resource Management *(US training course)*
ADP	Aéroports de Paris *(France)*
ALL	Adult Literacy and Lifeskills (survey)
AMAN	Arrival Manager
AOI	Area Of Interest
ATC	Air Traffic Control
ATCO	Air Traffic Controller / Air Traffic Control Officer *(US/UK)*
ATM	Air Traffic Management
ATSCQ	Air Traffic Control Safety Questionnaire
CAST	Consequences of future ATM systems for air traffic controller Selection and Training *(NLR)*
CCR	Central Control Room
CDM	Collaborative Decision-Making
CDMMA	CDM using Multiple Agent
CEC	Commission of the European Communities
CESC	Central Emergency Support Centre
CFMU	Central Flow Management Unit *(EUROCONTROL)*
CRDA	Converging Runway Display Aid
CRM	Crew Resource Management
CTA	Cognitive Task Analysis

DAP	Director(ate) ATM Programmes *(EUROCONTROL Headquarters, SD)*
DAS	Director(ate) ATM Strategies *(EUROCONTROL Headquarters, SD)*
DAS/HUM *or just* HUM	Human Factors Management Business Division *(EUROCONTROL Headquarters, SD, DAS)*
DG-TREN	Directorate General for Transport and Energy *(EU)*
DIS	Director(ate) Infrastructure, ATC Systems and Support *(EUROCONTROL Headquarters, SDE)*
DIS/HUM *or just* HUM	Human Factors and Manpower Unit *(EUROCONTROL Headquarters, SDE; formerly stood for 'ATM Human Resources Unit'; now known as 'DAS/HUM' or just 'HUM')*
DLR	Deutsches Zentrum für Luft und Raumfahrt *(German Aerospace Centre)*
DMAN	Departure Manager
DNA	Direction de la Navigation Aérienne *(Ministère de l'Equipement, des Transports et du Logement, France)*
EATCHIP	European Air Traffic Control Harmonisation and Integration *(renamed 'EATMP' in February 1999 and known as 'EATM' since May 2003)*
EATM(P)	European Air Traffic Management (Programme) *(known as 'EATCHIP' until February 1999)*
ECC	Emergency Control Centre
EEG	Electroencephalogram
EFS	Electronic Flight Strip
EMC[2]	Emergency Management Command and Control
EPE AENA	Entidad Pública Empresarial Aeropuertos Españoles y Navegación Aérea *(Spain)*
EPOG	Eye Point-of-Gaze
ER	Emergency Room

ESSAI	Enhanced Safety through Situation Awareness Integration in training
ET	Executive Task *(EATCHIP)*
EU	European Union
FAA	Federal Aviation Administration *(US)*
FMAQ	Flight Management Attitude Questionnaire
FPS	Flight Progress Strip
GAS	Gate Allocation System
GIHRE	Group Interactions in High Risk Environments
GUI	Guidelines *(EATCHIP/EATM(P))*
HFFG	Human Factors Focus Group *(EATM, HRT; formerly known as 'HFSG')*
HFSG	Human Factors Sub-Group *(EATCHIP/EATMP, HRT; today known as 'HFFG')*
HIPS	Highly Interactive Problem Solver
HRS	Human Resources Programme *(EATM(P))*
HRT	Human Resources Team *(EATCHIP/EATM(P))*
HUM	Human Resources (Domain) *(EATCHIP/EATMP)*
INECO	INgeniería y ECOnomía del Transporte *(Spain)*
ISDEFE	Ingeniería de Sistemas para la DEFensa de España *(Spain)*
ITA	Integrated Task Analysis
LEONARDO	Linking Existing ON-ground ARrival and Departure Operations *(EU Project)*
LOA	Level Of Automation
LOSA	Line Operation Safety Audit
MATT	Models for the Analysis of Team Training
ME	Mental Effort
MIL	Military

NASA	National Aeronautics and Space Administration *(US)*
NASA-TLX	Task Load Index - developed at NASA
NLR	Nationaal Lucht- en Ruimtevaartlaboratorium *(The Netherlands Aerospace Laboratory)*
NOTECHS	*Assessment method developed by DLR, IMASSA, NLR and University of Aberdeen*
NPP	Nuclear Power Plant
NRC	National Research Council *(US)*
NTS	Non-Technical Skills
NUTEC	Norwegian Underwater Technology Centre AS
OCTO	Offshore Command Training Organization
ODID	Operational Display and Input Development
OIM	Offshore Oil Installation Manager
OPITO	Offshore Petroleum Industry Training Organisation
ORMAQ	Operating Room Management Attitude Questionnaire
PAT	PHARE Advanced Tool
PC	Planning Controller
PD	PHARE Demonstration
PHARE	Programme for Harmonised Air Traffic Management Research in EUROCONTROL
PSG	Programme Steering Group *(EATM(P)/HRS)*
PVD	Plan View Display
R/T	Radiotelephony
REP	Report *(EATCHIP/EATM(P))*
RSME	Rating Scale ME
SA	Situation Awareness

SAGs	Severe Accident Guidelines
SBERGs	Symptom-based Emergency Response Guidelines
SD	Senior Director, EATM Service Business Unit *(EUROCONTROL Headquarters; formerly known as 'SDE')*
SDE	Senior Director, Principal EATMP Directorate *or, in short,* Senior Director(ate) EATMP (*EUROCONTROL Headquarters; now known as 'SD')*
SHAPE (Project)	Solutions for Human-Automation Partnerships in European ATM (Project) *(EATM(P), HRS, HSP)*
SASHA_Q	SHAPE SA Questionnaire
SKAs	Skills, Knowledge and Attitudes
SKATE	SKA in TEamwork (Model)
SICTA	Sistemi Innovativi per il Controllo del Traffico Aereo *(Italy)*
SME	Subject Matter Expert
SOIs	Station Operating Instructions
SPSS	Statistical Package for the Social Sciences
ST	Specialist Task *(EATCHIP)*
TC	Tactical Controller
TDB	Track Data Block
TDG	Tactical Decision Games
TIP	Taxi-In Planner
TMA	Terminal Control Area
TOP	Taxi-Out Planner
TP	Team Process questionnaire
TPF	Team Process Frequency
TPQ	Team Process Quality (rating) form

TRM	Team Resource Management
TRS	Tower Research Simulator
TSA	Team Situation Awareness
USQ	Uncertainty Sources Questionnaire
VINTHEC	Visual INTeraction and Human Effectiveness in the Cockpit

CONTRIBUTORS

The contribution of the Members of the HRT Human Factors Sub-Group (HFSG), now known as the Human Factors Focus Group (HFFG), to this document during the group meetings, and further written comments, were much appreciated.

Special thanks are also given to the QinetiQ Team.

Document configuration

C. HELLINCKX
(external contractor)

EUROCONTROL Headquarters

Page intentionally left blank

APPENDIX A DESCRIPTION OF OTHER DOMAINS

1. Nuclear Power Industry

1.1 Introduction

In the UK nuclear power industry 'accident management' corresponds to the same concept as 'emergency response', 'threat management' or 'crisis management'. Accident management refers to the set of actions that plant operators can take in order to cope with abnormal events, specifically by controlling the outcome of the event by intervening as early as possible.

A typical emergency response organisation does not consist of one dedicated team, but of several teams. These teams are embedded within the larger group of the entire emergency response organisation. The teams are located both on and off site. The on-site teams are at the Access Control Point, (ACP), the Central Control Room (CCR) and the Emergency Control Centre (ECC). The Central Emergency Support Centre (CESC) is located off site. Each team includes between two and eighteen persons. Some of these teams work together on a daily basis, other teams come together only in the case of an emergency.

Only recently the UK Power Industry has started paying attention to the Non-Technical Skills (NTS) necessary in accident management (Crichton & Flin, 1999). These researchers investigated what NTS are necessary to deal with emergency situations in nuclear power industry. Drawing on the good results found with Crew Resources Management training in aviation, they argue that technical skills are prerequisite, but not sufficient, to deal with all emergency situations. The goal of their study was to identify training for those skills in order to increase teamwork in responding to any emergency situation. Crichton and Flin analysed the roles and tasks of the different members of the teams and groups in the emergency response organisation at two UK power plants. The generic NTS that they found were communication, teamwork and stress management. Role-specific NTS included decision-making, leadership and situation awareness.

1.2 Avoiding accidents

The intention of nuclear power plant accident management is to gain control of the outcome of an abnormal event at the earliest possible time in order to prevent any adverse consequences. An accident condition is defined as the deviation from any normal state. A problem in identifying whether a state is abnormal is the (in) ability of designers of power plant to reliably predict the state of the plant *after* an abnormal event (e.g. fault). Therefore, it is not possible to counter all accident events with pre-defined procedures. As some accidents are caused by a combination of initiating (rare) events, Symptom-based Emergency Response Guidelines (SBERGs) have been introduced to support the power plant personnel in coping with accidents that have not been described by technical knowledge or experience. SBERGs supply general support to the operators based on symptoms only. This means that no diagnosis is necessary, in order to initiate actions to gain control of the accident.

1.3 Recognising threats

Three categories of non-normal states are distinguished. For each of these categories, different coping strategies have been developed. The non-normal situations that have been anticipated previously are identified by specific and recognised faults. In these situations, plant operators make use of Station Operating Instructions (SOIs). In the nuclear power industry it is acknowledged that not all possible faults and combinations of fault can be predicted and anticipated beforehand. Certain situations are characterised by a number of symptoms, but no specific diagnosis can be made based on technical knowledge. For such situations, Symptom-based Emergency Response Guidelines (SBERGs) have been devised.

SBERGs are developed to cover infrequent and ‘not credible’ faults. They offer general guidelines and advice (the advice being as accurate as possible) about what to do in order to prevent the fault from developing into a more serious situation. The focus of the guidelines is on the *symptoms* of the fault, and not the *diagnosis*. Instead of rigid adherence to procedures (SOIs), operators are encouraged and advised to make decisions regarding the appropriateness of any particular procedure. These decisions should be based on the information and resources available to them and the state of the plant. Finally, Severe Accident Guidelines (SAGs) cover situations that are recognised to be potentially threatening to the larger civil public. In these situations it is not advantageous to further analyse the symptoms of the non-normal events, but action should be taken to shut down (parts of) the plant.

In order to use the different strategies effectively, it is very important that the operators *recognise* the nature of the situation. Situation awareness is therefore critical in coping with the different sort of abnormal situations that can arise.

1.4 Team coordination

Different styles of decision-making have been identified, depending on the level of command, time pressure and available information. Crichton and Flin (1999) found two levels of decision-making – distributed and individual. Distributed decision-making, or coordination, takes place within teams. Distributed decision-making is determined a.o. by the individuals within the team. The types of individual decision-making corresponds to Rasmussen’s Skill- Rule- and Knowledge-based decision-making. The authors found that technical decisions were mainly made on the skill- and rule-based levels. The decisions made by team leaders ranged from analytic decision involving choices between alternative courses of action to fast, recognition primed intuitive decision. As team leaders have to deal with the analysis with the whole plant and the development of the situation, they need to make decisions on a knowledge-based level more than decision makers at a lower hierarchical level who make more technical decisions.

1.5 Performance of action

Most emphasis in research on emergency management is placed on the 'early' stages, i.e. situation assessment and awareness and the non-technical skills necessary in these stages. Much less attention is paid to the later stages, e.g. performance of actions after the situation assessment has been completed. This can be explained by arguing that NTS play a minor role in the 'acting' stages of emergency management: orders have been given by the higher levels in hierarchy, and the technical personnel simply carries out these orders (action). Next to that, little is said about monitoring the results of the actions that are carried out.

1.6 Communication

Communication responsibility depends for a large part on the role in the team of a person: key players are heavily burdened by their responsibilities (e.g. communications between on- and off-site). Personnel with other roles (e.g. off-site technical assistance) are less dependent and responsible for accurate and flexible communication. This is due to delays in information supply: as a result there is limited opportunity to provide suggestions or advice. Particular communication skills can be distinguished: briefing, advising and listening, as well as non-verbal communication.

1.7 Training of non-technical skills

Crichton and Flin (1999) argue that training intervention should aim for the improvement of the Non-Technical Skills (NTS). Training should be tailored to the specific roles in nuclear power plants. Specifically, Crichton and Flin address Crew Resource Management (CRM) and Tactical Decision Games (TDGs) workbooks. They state that these types of training are useful for teams that work with each other on a day-to-day basis, but especially effective for teams that only come together in response to an emergency. These training interventions should aim for the development of rules of thumb that are applicable in a repertoire of incidents. CRM-training results in two improvements. Firstly, CRM improves teamwork in order to minimise the risk of occurrence of emergencies or accidents. Secondly, CRM helps crew in dealing with emergencies more effectively once the emergency has occurred.

TDGs are relatively novel training methods. The goal through these TDGs is to exercise decision-making skills and illustrate key operating principles. More importantly, TDGs should enable the participants to develop a shared understanding and recognition of possible problems for emergency management. TDGs offer a variety of scenarios differing in complexity, timescale and technicalities.

2. Medical Domain (Anaesthesia)

2.1 Introduction

The investigation was restricted to medical teamwork and decision-making as it takes place in the operating theatre at a hospital. More specifically, teamwork inside this theatre (i.c. an anaesthetic team) was investigated. A core anaesthetic team consists of one anaesthetist and one anaesthetic nurse, but can be extended with extra staff of either type as special circumstances require. The role of the anaesthetist in the operating theatre is essentially threefold. Firstly to ensure that the patient is unaware of and cannot recall any of the surgical stimuli. Secondly, to provide conditions that will allow the surgeon to carry out the operation. Thirdly to ensure that the patient experiences as little morbidity as possible.

At first glance anaesthesia might seem to bear little resemblance with ATC work. In fact, this domain shares a number of psychological similarities with ATC work, as far as sequential phases in operational work with various degrees of workload can be distinguished and vigilance, monitoring, situation assessment, decision-making and team coordination are concerned.

In anaesthetics, like in ATC, continuously more trust is placed on technology. Elder anaesthetists used to do small repair jobs when their equipment had certain failures. Nowadays, only the 'self-test function' of digital equipment can be operated in case of (suspected) malfunction. Basic standard equipment consist of anaesthetic apparatus, consisting of (electronic) breathing apparatus and monitoring apparatus, checking heart, blood pressure, oxygen saturation, CO_2 and temperature. Other standard apparatus include an infusion apparatus, pumps for medication and heat mattresses. Less standard equipment is a bi-spectral analysis apparatus, to measure the depth of the anaesthesia, via ElectroEncephaloGram (EEG).

At the regional hospital visited, ca. 20% of all operations are emergency operations, and ca. 80% are elective operations. Emergency operations are those that come in through the Emergency Room (ER) and need to be operated on within 24 hours. In such cases it is less possible to prepare for the operation. Within this group 5% of the patients needs to be operated upon within half an hour. This happens about once in every two days for an anaesthetist.

Although preparation (e.g. through familiarisation with the patient) is not possible in the 20% group, it is still considered as standard work. Uncertainty, e.g. on the patients reactions, is increased, but coping with this type of uncertainty is still considered as part of the standard job. Within this group morbidity and mortality are higher than in the 80% group. This is because 1) usually, the patients are simply in a worse condition when they arrive than elective patients and 2) the uncertainty during operation increases the risks involved.

In the earlier mentioned 20% group of patients, it is usually necessary to be extra alert for sudden mishaps, such as sudden blood loss. Although there is little or no time to do pre-operative assessments on the patient, special preparations can be made in advance, e.g. to make sure that extra blood can be administered on the shortest timeframe possible. Again, all situations described are fairly standard, and the difficulty is that one does not know when they will happen.

2.2 Communication

Anaesthetists are team players with a tendency to do things their way. Although they receive years of training and practice for decades, their syllabuses rarely include communication techniques. According to the interviewee, an important channel of communication is the status report of the patient, which is used by all treating doctors. Sometimes the handwriting is unreadable, or things are not written down at all, or ambiguous in their interpretation. For example, one doctor might think that ‘intubating’ a particular patient proved very tricky, whilst another doctor might not find it difficult at all. In those status reports there is some sort coding for the ‘difficulty’ of the patient (e.g. how difficult the infusion was). 1-4, 1 = easy, 4 = almost impossible. There is also a ‘protocol’ that all decisions etc. are written down and passed along to the other professionals that have anything to do with it. This is to trace back all steps if necessary. During surgery, obviously there is no time to record everything. As stated before, a lot is written down on the status report of the patient, however, because of the direct contact in the operation theatre, most communication is oral. This is through ‘free speech’, no read-back procedures or confirmations are used. Two different sources of anaesthetic incident data referred to by Fletcher *et al.* (2002) reveal that poor communication is a contributory factor in respectively 34% and 9% of the incidents. To illustrate, causal factors for poor communication in a particular class of anaesthetic problems (‘airway management’, i.c. airflow through throat, trachea, etc.) that were mentioned:

- knowledge gaps (e.g. an incomplete understanding of the implications of certain diagnostic signs or the use of infrequently used devices);
- inappropriate fixation on a problem;
- excessive discussion about non-task relevant issues while failing to monitor patient vital signs that would have indicated de developing complication;
- failure to follow established procedures, particularly those involving evaluation and re-evaluation of patient status;
- ineffective team coordination and use of available personnel resources;
- a lack of preparation and checking of instrumentation prior to use.

2.3 Training of non-technical skills

To become a qualified anaesthetist the candidate receives on-the-job training for seven years (after being qualified as an MD). During this period about twenty different experienced anaesthetists supervise the student. No further specific NTS training is received.

Most training and training material in anaesthesia primarily deals with the patient's normal or abnormal physiology, or with the technical and clinical characteristics of drugs and equipment. However, over the last decade, a slight shift of focus has taken places toward the NTS of the anaesthetist, that is how to understand, prepare for, and manage crisis situations.

From the literature review by Fletcher *et al.* (2002) it was learned that in the US Anaesthesia Crisis Resource Management (ACRM) training was devised to help address the problem of limited exposure to emergency situations. Following on from CRM training in aviation a course was developed for anaesthetists to help them develop ‘pre-compiled responses to critical incidents’ and ‘to instruct participants in the coordinated integration of all available resources to maximise safe patient outcomes’. The course involved four main stages plus completion of evaluation questionnaires. The first stage was an introduction to ACRM through lectures and discussions on human performance and human factor topics, videotapes showing a simulated aircraft accident and an actual anaesthetic incident. The final teaching element was discussion of teamwork (in managing the emergency) from the videos and introduction to a set of CRM principles. The next stage of the course was for familiarisation with a full physical and functional simulation of the operating theatre (including a body mannequin, computer simulations and standard equipment located in a normal operating theatre), followed by actual training sessions. These consisted of a number of different scenarios each offering the opportunity for an individual (doctors and nurses) to be in the ‘hot seat’. The final stage of the training involved debriefing with the session video. Post-evaluation of the course, showed that the course was valuable, in particular the simulator session. The taught component of the course on Human Factors and Human Performance was least popular, even though this section was aimed to raise understanding of the concept as a whole.

It is characteristic that a current obstacle for assessing the effectiveness of ACRM training is the absence of a validated performance measure for resource management behaviours.

3. Offshore Oil Industry

3.1 Introduction

An important cause of the attention for emergency decision-making in this domain was the disaster with the Piper Alpha. We use this disaster to sketch the circumstances of Crisis Management in this domain. The Piper Alpha oil and gas production platform in the North Sea located 110 miles Northeast of Aberdeen. At around 22:00 on the 6th of July 1988 there was an explosion on the production deck of the platform, which was caused by the ignition of a cloud of gas leaking from a temporary flange. The resulting fire spread rapidly and was followed by a number of smaller explosions. At around 22:20, there was a major explosion caused by the rupture of a pipeline carrying gas to the Piper Alpha from a nearby platform. Over the course of the next few hours an intense high pressure gas fire raged, punctuated by a series of major explosions that served to hasten the structural collapse of the platform. Of the 226 persons who were on board the installation, only 61 survived. The great majority of the survivors escaped by jumping into the sea, some from as high as 175 feet. It appears that the crisis on board of the Piper Alpha could have been managed more effectively. The official enquiry contained a number of criticisms relating to the performance of Piper Alpha's Offshore Oil Installation Manager (OIM) on the night of the disaster as well as the OIMs on duty on the adjacent platforms. These installations were linked to the Piper Alpha by hydrocarbon pipelines, the rupturing of which caused massive explosions and the rapid spread of the fire on the Piper Alpha. It was suggested that, had the production of hydrocarbons by the adjacent platforms been stopped sooner, the situation on the Piper Alpha might have deteriorated less rapidly.

3.2 Avoiding threats

Following the Piper Alpha disaster and the enquiry, various recommendations of the official enquiry were implemented at offshore companies, such as:

- The creation of an personnel assessment process that would maximize the learning opportunities for individuals yet would be constructive and supportive. The process could not be based on a pass/fail regime similar to the assessment used by the Royal Navy for its submarine commanders, where failure (historically around 25%) marks the end of an officer's career in submarines. However, similar elements of such assessment were considered by a major offshore oil company to be entirely appropriate. In addition, it was intended to develop emergency handling skills, build teams that would be capable of supporting the OIM even in the worst scenarios, and create a system that was not centred on a few charismatic personnel.

- Intensive training in emergency command. One of the newest simulators is that of the Norwegian Underwater Technology Centre AS (NUTEC) in Bergen, Norway, which trains very large offshore teams in a purpose built facility. This is based in a suite of rooms with full communication links, recording equipment and computer controlled scenarios, in which not only the platform emergency response team participate, but also helicopter and ship captains who would be involved in a rescue operation. The simulator is built on the side of a fjord, and when the platform manager takes the decision to evacuate installation, the crew have to leave simulator and board a free-fall lifeboat, which they then have to launch. The latter component injects a degree of realism found in few other industrial command simulators.

3.3 Preparing to manage the threat

As was found also in other domains (e.g. medical decision-making), commanders who have thought through (mentally rehearsed) their emergency response plan should find it easier to implement these actions when operating under stress. One offshore manager said that his advice to an inexperienced manager who had to deal with an offshore emergency would be, ‘keep calm or you lose it straight away. You must have contingency plans in your own mind, that is thought out possible scenarios and responses before incidents occur’ (Flin & Slaven, 1994). So, while the prime function of standard response procedures is to assist in decision-making and team coordination, it should be noted that they may also help to mitigate the effects of stress for those in command positions.

3.4 Team coordination and action planning

An Offshore Oil Installation Manager (OIM) reports to more senior managers who are based onshore. In the event of a serious offshore incident, an onshore emergency response team is gathered in the company’s shore-based office, some of which have dedicated emergency control centres. This can create a distributed decision-making situation, with onshore managers and specialists becoming involved. For some OIMs this had been helpful in preventing media interference, organising evacuation helicopters, providing specialist knowledge and dealing with relatives’ enquiries. Where the OIMs felt that the onshore/offshore interface caused problems was when the onshore managers failed to realize the severity of the situation, questioned the OIM’s decisions, needlessly rearranged evacuation arrangements resulting in serious confusion, asked for information rather than giving it, or did not know enough about the installation.

An interesting point was made on risk and time assessment. According to Prof. Flin (see also Flin, 1996), several OIMs commented on the rapid pace of decision-making required during the incident: ‘I probably made hundred decisions in a short space of time’. The time pressure variable is elemental in the natural decision-making situation and the decision maker’s judgement of the time horizon is critical. According to Orasanu and Fischer (1997) *situation assessment requires definition of the problem and assessment of risk level and time available to make the decision*. The OIMs felt that it was vitally important to be decisive in what might be a rapidly escalating situation. They also appreciated that in some instances, because of the time pressure, they would have to take a satisfactory rather than an optimal decision. In the words of one OIM 'minutely wrong decisions are better than no decisions at all'. The decision-maker's awareness that the aim is to reach a satisfactory rather than an optimal solution is characteristic of experts using recognition-primed decision-making (Klein, 1989).

3.5 Communication

According to the standard of competence 'controlling emergencies' for offshore managers of the Offshore Petroleum Industry Training Organisation (OPITO), the maintenance of communication entails:

- all essential people and organisations are immediately informed of the emergency;
- reports of the situation as it develops are provided to installation staff at suitable intervals;
- appropriate communications are maintained during the emergency;
- an accurate record of all events and of key communications is maintained.

In a CRM course for offshore control room operators' emergency response training, a communication module covered the basic communication process, barriers to effective communication and awareness of strengths and weaknesses in personal communication skills. Exercises highlighted the importance of feedback and listening skills, the role of non-verbal communication and effective communication techniques. An actual offshore incident involving a communication problem was also presented and discussed.

3.6 CRM and threat management courses

A number of organisations now offer specialist courses, such as Emergency Management Command and Control (EMC^2) or Offshore Command Training Organization (OCTO). These courses draw on the expertise of former Naval commanders (including, in the latter case, two admirals from the Falklands conflict) to provide intensive coaching and development of managers, using both outdoor training and emergency command centre simulator exercises. The oil companies also conduct regular training on their installations, from weekly lifeboat drills to full scale exercises.

One of the conclusions of the article by Flin, Slaven and Stewart (1996) was that although OIMs exhibited decisions based on recognition of the situation, it does not seem effective to train for recognition primed decision-making. Recognition is a skill that develops through experience, and therefore cannot be trained directly. However, training should focus on situation awareness, searching and understanding critical cues available to the OIMs. Next to that, training should focus on the accuracy of the risk and time judgements that have to be made by the OIMs.

Page intentionally left blank

APPENDIX B QUESTIONNAIRES APPLIED

Questionnaire

To be completed by each team member at the end of the experiment

Date: Time: ID:	Your (most frequent) controller role during the trials: Runway Controller ❑ Departure/Start-up Controller ❑ Ground Controller ❑

1. Please read carefully through the list of statements.
2. All statements relate to teamwork in the Phase II trials, that is, with the Collaborative Decision-Making (CDM) concept/system.
3. Indicate to which extent you agree with this statement by putting a cross on a scale from 1 (strongly disagree) to 5 (strongly agree).
4. When the statements make a comparison (for example: 'CDM reduces the need for team communication'), think of the reference system (Phase I) as the basis for this comparison.
5. Our definition of the team is: Runway Controller, Departure/Start-up Controller and Ground Controller.
6. Complete the questions from the viewpoint of the controller role you had most frequently during the trials.
7. You can give remarks or suggestions at the end of this form. Please provide the number of the statement to which your remark or suggestion relates.

	Disagree Strongly	**Disagree Slightly**	**Neutral**	**Agree Slightly**	**Agree Strongly**
1. With the CDM system, it is clear to me which tasks are my own responsibility and have to be done manually.	1	2	3	4	5
2. With the CDM system, it is clear to me which tasks are done entirely by the system.	1	2	3	4	5
3. With the CDM system, it is clear to me which tasks are done entirely by the other team members, with or without the system.	1	2	3	4	5

	Disagree Strongly	Disagree Slightly	Neutral	Agree Slightly	Agree Strongly
4. With the CDM system, it is clear to me which tasks I share with the other team members.	1	2	3	4	5
5. The CDM system allows the team to prioritise tasks more efficiently.	1	2	3	4	5
6. With CDM, there is less need to ask other team members for assistance.	1	2	3	4	5
7. With CDM, there is less need to discuss working strategies with other team members.	1	2	3	4	5
8. My working style is compatible with the CDM system.	1	2	3	4	5
9. With the CDM system, my working style is no longer compatible with that of the other team members.	1	2	3	4	5
10. CDM reduces the need for team communication.	1	2	3	4	5
11. CDM promotes team spirit.	1	2	3	4	5
12. With CDM, teamwork becomes more important.	1	2	3	4	5
13. I have more trust in my team members when using the CDM system.	1	2	3	4	5
14. With CDM, the goals of the team are more clearly defined.	1	2	3	4	5
15. The team has the appropriate size for working with the CDM system.	1	2	3	4	5
16. With CDM, the responsibilities of the team members are clearly defined.	1	2	3	4	5

	Disagree Strongly	Disagree Slightly	Neutral	Agree Slightly	Agree Strongly
17. The team continues to learn to cooperate in working with the CDM system.	1	2	3	4	5
18. The CDM system promotes clarity of roles in the team.	1	2	3	4	5
19. The CDM system helps me to detect other team members' inaccuracies or mistakes.	1	2	3	4	5
20. The CDM system helps me to take initiative if required by the situation.	1	2	3	4	5
21. The CDM system helps me to take command if required by the situation.	1	2	3	4	5
22. The CDM system helps me to state my plans when required.	1	2	3	4	5
23. The CDM system helps me to change plan, with other team members being informed/consulted.	1	2	3	4	5
24. With CDM, the task distribution among the team is appropriate.	1	2	3	4	5
25. The CDM system helps me to intervene if task completion deviates from standards.	1	2	3	4	5
26. The CDM system encourages the team to monitor each other.	1	2	3	4	5
27. The CDM system helps me sharing information about developing traffic situations with others.	1	2	3	4	5

	Disagree Strongly	Disagree Slightly	Neutral	Agree Slightly	Agree Strongly
28. The CDM system helps team members alert each other to new events.	1	2	3	4	5
29. The CDM system helps the team to synchronize their actions so as not to hinder or delay each other.	1	2	3	4	5
30. The CDM system makes it easier for the team to identify future overloading.	1	2	3	4	5
31. It is harder to motivate the team when working with the CDM system.	1	2	3	4	5
32. When working with the CDM system, none of the team members needs to take command over the others.	1	2	3	4	5
33. With CDM, coordination between team members remains necessary.	1	2	3	4	5

If you have any additional comments, please add them here:

This is the end of the questionnaire, thank you for your cooperation!

Self-rating Form

To be completed by each team member after each run

Debriefing date: ______________________ Run ID: ______________________ Scenario: Phase I (reference) ❑ Phase II (CDM) ❑	Your function in the team during last run: Runway Controller ❑ Departure/Start-up Controller ❑ Ground Controller ❑

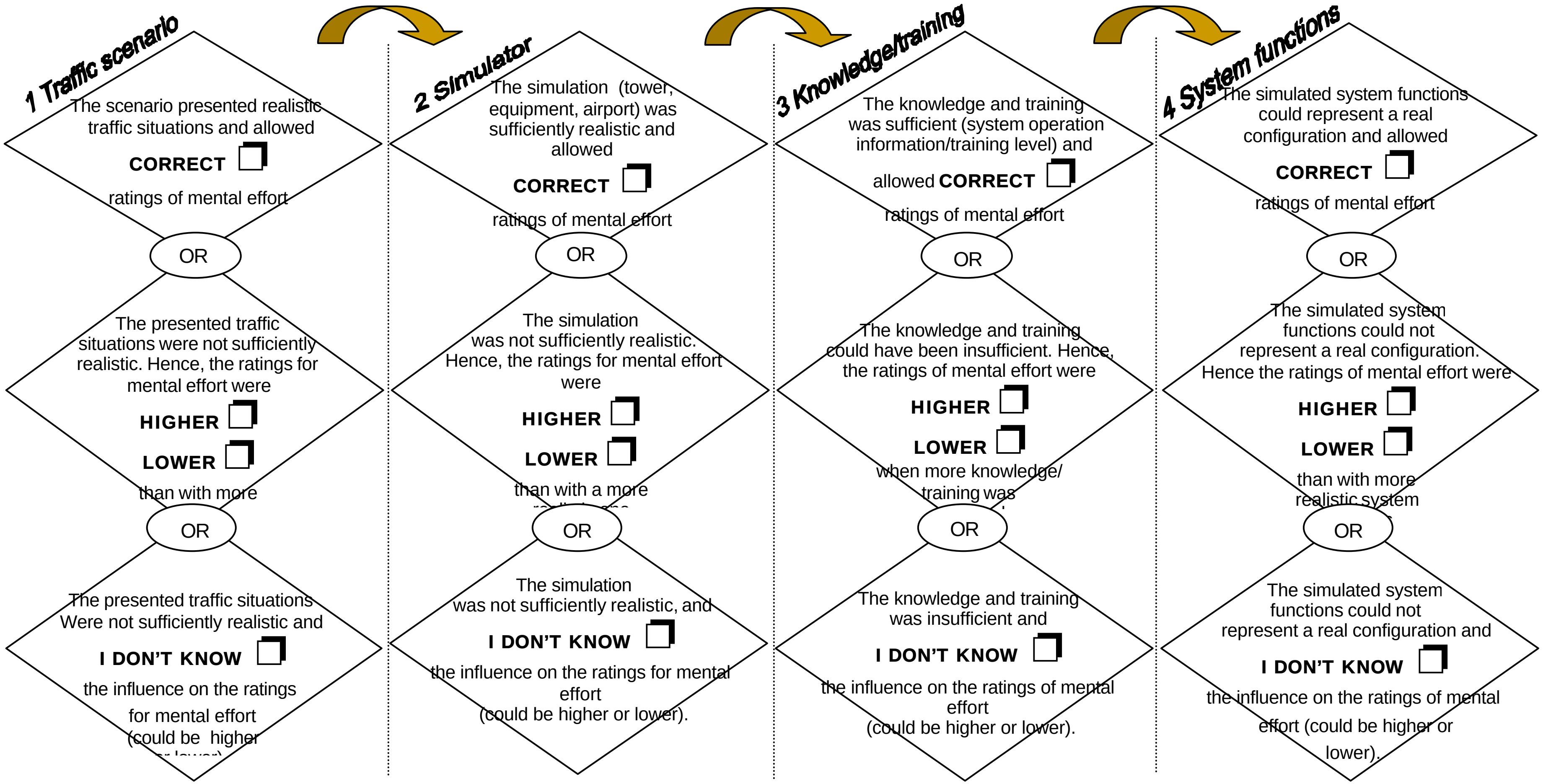
Mental Effort of Team Members

Please indicate, by placing a mark on each of the vertical lines below, how much mental effort each team member (including yourself) had to spend in order to cope with the events in the scenario. Try to make estimates for the other team members. Do not ask or negotiate these estimates with the other team members.

Runway Controller	***Departure/ Start-up Controller***	***Ground Controller***
15	15	15
14	14	14
13	13	13
12	12	12
Tremendously effortful	Tremendously effortful	Tremendously effortful
11	11	11
Very, very effortful	Very, very effortful	Very, very effortful
10	10	10
9	9	9
Very effortful	Very effortful	Very effortful
8	8	8
7	7	7
Pretty effortful	Pretty effortful	Pretty effortful
6	6	6
Rather effortful	Rather effortful	Rather effortful
5	5	5
4	4	4
Somewhat effortful	Somewhat effortful	Somewhat effortful
3	3	3
A bit effortful	A bit effortful	A bit effortful
2	2	2
1	1	1
Hardly effortful	Hardly effortful	Hardly effortful
Not at all effortful	Not at all effortful	Not at all effortful
0	0	0

Uncertainty Sources in Ratings of Mental Effort

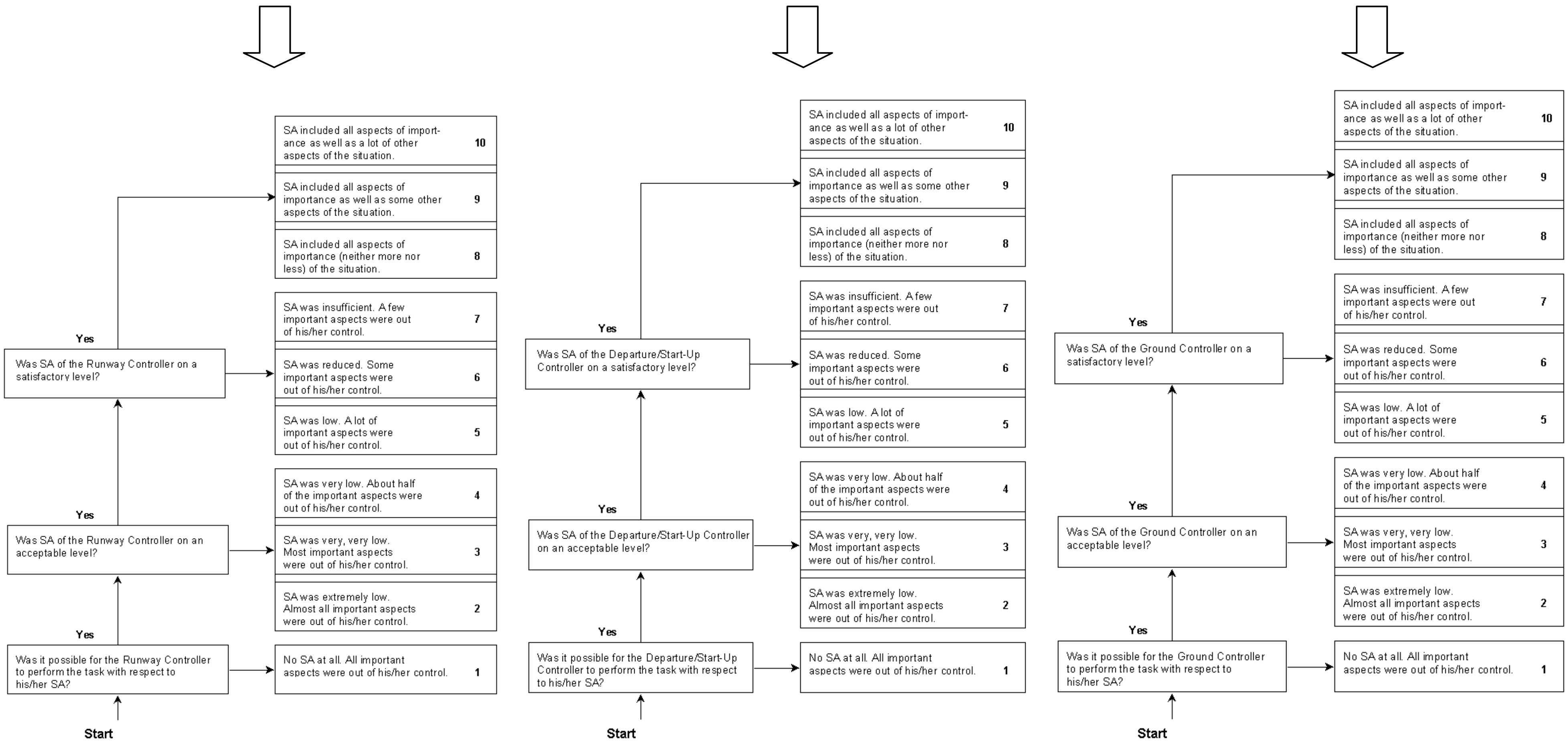
The goal of the diagram is to evaluate the effect of various aspects of the exercise on the ratings for mental effort that you just provided. The quality of these aspects (which are: *traffic scenario*, *simulator*, *knowledge/training provided* and *system functions*) could have lead to different ratings of mental effort than those given in a real situation. For each aspect you put one cross at either 'correct', 'higher', 'lower' or 'I don't know'.



Situation Awareness (SA) of Team Members

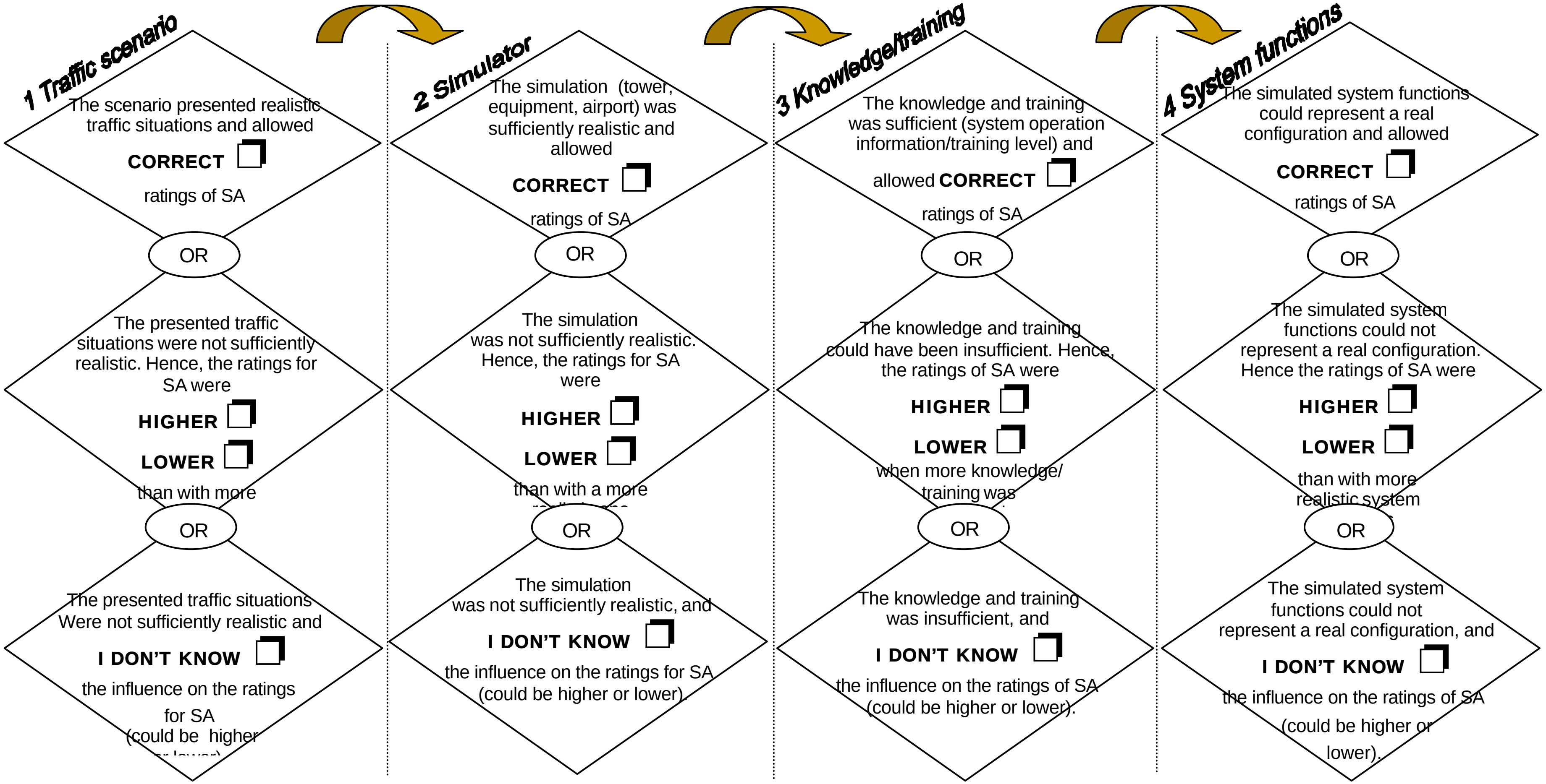
Please indicate the level of Situation Awareness (SA) for each of the team members (including yourself) during the last run. Encircle the appropriate number (one to ten) for each team member.

Runway Controller **Departure/Start-up Controller** **Ground Controller**



Uncertainty Sources in Ratings of Situation Awareness

The goal of the diagram is to evaluate the effect of various aspects of the exercise on the ratings for SA. The quality of these aspects (traffic scenario, simulator, knowledge/training provided and system functions) could have lead to different ratings of SA than those given in a real situation. For each aspect you put one cross at either 'correct', 'higher', 'lower' or 'I don't know'.



Additional question on your Situation Awareness (SA). Please answer each question by ticking the box as appropriate.

	Never				Always
Q1. Did you have the feeling that you were ahead of the traffic, able to predict the evolution of the traffic?	1	2	3	4	5
Q2. Did you have the feeling that you were able to plan and organise your work as you wanted?	1	2	3	4	5
Q3. Have you been surprised by an aircraft call that you were not expecting?	1	2	3	4	5
Q4. Did you have the feeling of starting to focus too much on a single problem?	1	2	3	4	5
Q5. Did you forget to act on any aircraft?	1	2	3	4	5
Q6. Did you have any difficulty finding an item of (static) information?	1	2	3	4	5
Q7. Do you think that the system provided you with useful information?	1	2	3	4	5
Q8. Were you paying too much attention to the functioning of the system?	1	2	3	4	5
Q9. Did the system help you to have a better understanding of the situation?	1	2	3	4	5

	Poor	**Quite poor**	**Okay**	**Quite good**	**Very good**
Q10. Finally, how would you rate your overall situation awareness during the exercise?	1	2	3	4	5

If you have any additional comments, please add them here:

This is the end of the questionnaire. Thank you for your cooperation!

APPENDIX C OBSERVATION FORM

Team ID:	Observer:	Start time recording:	Duration of interruptions:
Phase I /Phase II	Date:	End time recording:	Duration or recording:
Depart. rwy:	Arrival rwy:		

	Team Members					
Communication **Type & Content**	**Departure Ctrl / SUC**		**Tower Controller**		**Ground Ctrl**	
	® **TWR Ctrl**	® ® **Gnd Ctrl**	¬ **Dep/SUC**	® **Gnd Ctrl**	¬ **TWR Ctrl**	¬ ¬ **Dep/SUC**
Asks for ***Information***						
Asks someone to perform ***Action***						
Other requests						
Gives Information/ Answer						
States that he/she will initiate **Action**						
Suggests Action or **Solution**						
Other Announcements						
Acknowledgements Accept suggestion, solution or criticism. State that requested action will be provided.						

Page intentionally left blank

APPENDIX D DATA AND RESULTS

1. Overview of Results per Questionnaire (Q) Item

Below is the extended version of the tables for validation of the 33 questionnaire-(Q) items, already included in Section 7. The difference is that here the hypotheses tested in view of item validation are described and the results used for validation are summed up. References to a number of tables and figures can also be found, which provides the reader with the actual data upon which the validation is based.

1. With the CDM system, it is clear to me which tasks are my own responsibility and have to be done manually.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 5
Hypothesis	No difference in eye movements between the different interfaces was expected.
Summary of result	No difference found. Fixation frequency (per min) (Figure 19) Departure planner's Left + middle monitor for a whole process: R2CDMM2 < R3REFM2, R4CDMS2 < R6REFS2, no over all trend identified. Dwell frequency (per min) (Figure 20) demonstrated no consistent trend. Fixation frequency (per dwell) (Figure 21) Within middle monitor and right monitor: R2CDMM2 < R3REFM2, R4CDMS2 < R6REFS2. Fixation transition between different monitors (Figure 22). Transition from the middle monitor to the right monitor is more frequently in automation-on condition than automation-off condition (R2CDMM2 > R3REFM2, R4CDMS2 > R6REFS2), transition from the right monitor to the left monitor is less frequently in automation-on condition than automation-off condition, the transition from the right monitor to the middle monitor is more frequently in automation-on condition than automation-off condition.

2. With the CDM system, it is clear to me which tasks are done entirely by the system.	VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 2 Ground controller: 4
Hypothesis	In mixed RWYs scenario less gaze at arrival-departure table.
Summary of result	Average 48.5 % of fixation duration time looked at arrival-departure table in automation-off, 37.1 % of fixation duration time looked at arrival-departure table automation-on.

3. With the CDM system, it is clear to me which tasks are done entirely by the other team members, with or without the system.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 5 Ground controller: 5
Hypothesis	In the automation-on condition more time will be spend on the middle monitor (arrival-departure table) and less monitoring looks will be made at ground controller made, both by the departure planner.
Summary of result	With regard to the looking on middle monitor (arrival-departure table), that part of the hypothesis could be validated for mixed RWYs (see also item 2), but not for segregated RWYs (40.5% for automation-on, 32.9% for automation-off). The monitoring look (at ground controller) was observed less frequently in automation-on condition (0.6%) than in automation-off condition (6.7%) for mixed RWYs. However, for segregated RWYs automation-on condition shows more frequent monitoring look (0.5%) than automation-off condition (0.1%). Therefore, with regard to the monitoring look, validated for mixed RWYs, but not for segregated RWYs. Note: Fixation distribution over various AOIs (Figure 12, Figure 13), and monitoring look (Figure 14) have not provided consistent trends.

4. With the CDM system it is clear to me which tasks I share with the other team members.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 5 Ground controller: 4
Hypothesis	Less monitoring looks at other controllers and less communication in general between departure planner and others expected in automation-on mode.
Summary of result	Less monitoring looks validated for mixed mode, not for segregated mode (like item 3). There was very few communication taking place. However, on average, during automation-on 14.5 expressions were exchanged while during automation-off 29.8 expressions were exchanged. Note: Fixation distribution over various AOIs (Figure 12, Figure 13), and monitoring look (Figure 14) have not provided consistent trends.

5. The CDM system allows the team to prioritise tasks more efficiently.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 2 Ground controller: 5
Hypothesis	Slightly more eye movements made by the departure planner between the items on the screen (the task to prioritise is regarded slower) in automation-on condition.
Summary of result	The fixation frequency per dwell on the middle monitor (arrival-departure table) was lower in the automation-on condition compared to the automation-off condition. With regard to the left monitor (EFS), the same tendency was observed in segregated mode only. For the right monitor, this tendency was observed in mixed mode only. (Figure 21).

	The fixation transition between various monitors shows that transition from the middle monitor to the right monitor is more frequently in the automation-on condition than in the automation-off one (Figure 22). However, transition from the right monitor to the left monitor is less frequently in automation-on condition than automation-off condition, the transition from the right monitor to the middle monitor is more frequently in automation-on condition than automation-off condition. Over all, more transitions were made in the in automation-on mode.

6. With CDM, there is less need to ask other team members for assistance.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 5	
Hypothesis	The departure planner will ask more frequently for assistance in automation-on mode, while the ground controller will ask less frequently for assistance in automation-on mode.	
Summary of result	Not sufficient data. (There was very few communication taking place, especially when only the issue 'asking for assistance' is relevant. Therefore no statements about the trends regarding frequency of communication based upon these questionnaires are made.)	

7. With CDM, there is less need to discuss working strategies with other team members.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 5	
Hypothesis	For departure planner no change in communication frequency, or information exchange, in the observation form is expected. For the ground controller it is expected that the communication frequency, or information exchange, will decrease in the automation-on condition.	
Summary of result	On average over all runs the departure planner makes four verbal expressions in automation-on condition versus 0 in automation-off. The ground controller makes twice as much (13.5) verbal expressions in automation-off versus (6) automation-on condition.	

8. My working style is compatible with the CDM system.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4	
Hypothesis	It is expected that the randomness (entropy) in eye scanning pattern between automation-on and off condition will be the same.	
Summary of result	Because the samples turned out to be too small for appropriate entropy calculation and because the small differences would get lost in the bias, it turned out not to be possible to validate this item.	

9. With the CDM system, my working style is no longer compatible with that of the other team members.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 3
Hypothesis	The departure planner's eye scanning randomness is supposed to increase in the automation-on condition. The departure planner's differences in SA ratings for self and others are supposed to be smaller in automation-on condition.
Summary of result	Like described under item 8 entropy calculations were not applicable in the current study. However, the SA rating for self and other is about the same, or a very small difference between automation-on and off. Therefore this item is considered validated.

10. CDM reduces the need for team communication.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 5
Hypothesis	It is expected that the departure planner communicates more and looks more frequently at his colleagues in the automation-on condition.
Summary of result	The observation form does not confirm the hypothesis. Amongst others because very little communication took place. But the form shows that departure planners communication in general is more frequent in the automation-off condition (which is contradictory to the hypothesis).

11. CDM promotes team spirit.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 4
Hypothesis	Cannot be validated with the current experimental design.
Summary of result	Not validated.

12. With CDM, teamwork becomes more important.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 2
Hypothesis	This item could not be validated with the current data set. Therefore it was decided to interview the participants later about this matter.
Summary of result	The participating ATCOs have informed the researchers of the current study that teamwork will not become more difficult but definitely more important. Because more information will become available it will easier to inform and be informed. Teamwork will become more important, because decisions may have a direct impact on the process that the automation processes.

13. I have more trust in my team members when using the CDM system.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 3
Hypothesis	Departure planner will rate lower on SASHA question 8 (Were you paying too much attention to the functioning of the system?) with CDM on. For ground controller no difference in rating was expected. More monitoring of other team members by departure planner in automation-on mode.
Summary of result	There was no big difference in rating for departure planner as well as for ground controller, therefore validated for ground controller, not for departure planner. Note: Fixation distribution over various AOIs (Figure 12, Figure 13), and monitoring look (Figure 14) have not provided consistent trends.

14. With CDM, the goals of the team are more clearly defined.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 5
Hypothesis	Ground controller will have (rated by colleague) a higher SA in automation-on condition (compared to automation-off condition), and departure planner the other way around.
Summary of result	SAs for both are about equal in both conditions.

15. The team has the appropriate size for working with the CDM system.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 5 Ground controller: 4
Hypothesis	Higher SA rating expected for both in automation-on, compared to off, condition.
Summary of result	SAs for both are about equal in both conditions.

16. With CDM, the responsibilities of the team members are clearly defined.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 5 Ground controller: 5
Hypothesis	The departure planner will glance less frequently at the monitors of other team members, and will glance less frequently at his own monitor with the map of the Schiphol airport taxiways, in the automation-on condition. There will be less communication between the ATCOs in the automation-on condition.
Summary of result	With regard to segregated mode, the monitors of other team members were less frequently observed in automation-on condition than in automation-off condition. For mixed mode is not possible to verify this phenomenon. This indicates that the item can be validated. The map of the Schiphol airport was more frequently observed in

	automation-on condition than in the automation-off condition. This pleads against validating the item. Concerning communication, on average, during automation-on 14.5 expressions were exchanged while during automation-off 29.8 expressions were exchanged. So the second part of the hypothesis is validated.

17. The team continues to learn to cooperate in working with the CDM system.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 4 Ground controller: 3	
Hypothesis	In later experimental runs there will be less monitoring looks at other team members, or their monitors, and less communication exchange between team members.	
Summary of result	Figure 14 demonstrates that there is no consistent trend in either the monitoring of other ATCOs, or the monitoring of others ATCOs' monitors, by the departure planner in the different runs. Therefore this questionnaire item cannot be validated based upon the eye scanning behaviour of the departure planner. Amounts of communication exchanges per run: 1. 12 2. 9 3. 9 4. 5 5. 5 6. 29 The trend seems to be there, though the last run (6) forms a striking exception, so validation will not be based upon communication frequency solely.	

18. The CDM system promotes clarity of roles in the team.		VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 2 Ground controller: 5	
Hypothesis	With automation-on the departure planner will look slightly more often at the parts of the monitor that are not strictly necessary for his own task, but which will inform him about the status and work progress of the other team members.	
Summary of result	Fixation distribution over various AOIs is shown in Figure 12 and Figure 13. This distribution shows that gazes on the left and middle monitor by the departure planner were slightly less frequently observed in the automation-on condition than in the automation-off condition. (EPOG: automation-on < automation-off. Video-based analysis: R4CDMS2 < R6REFS2). With regard to the right monitor (map of Schiphol taxiways) it was the other way around. (EPOG: automation-on > automation-off, video-based analysis: R4CDMS2 > R6REFS2). The right monitor is considered a bit less elementary for the departure planner, compared to the other two monitors, as such this item is considered validated.	

19. The CDM system helps me to detect other team members' inaccuracies or mistakes.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 3
Hypothesis	Departure planner will spent more, or equal, amount of time in automation-on condition to monitor the slips and mistakes of team members.
Summary of result	Based on EPOG data, the monitoring of the ground controller by the departure planner was observed less frequently in the automation-on condition (0.6%) compared to the automation-off condition (6.7%) for the mixed RWYs scenarios. However, for segregated RWYs automation-on condition shows more frequent monitoring look (0.5%) than automation-off condition (0.1%). Therefore, with regard to the monitoring look, validated for mixed RWYs, but not for segregated RWYs. Two comparisons show different tendencies. Note: Fixation distribution over various AOIs (Figure 12, Figure 13), and monitoring look (Figure 14) have not provided consistent trends.

20. The CDM system helps me to take initiative if required by the situation.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4
Hypothesis	Less mental workload would enable the ATCOs to spend more capacity on taking initiative if required by the situation. The hypothesis is that there will be no difference in reported mental workload between automation-on and -off condition.
Summary of result	No difference in reported mental workload was expected, and no difference was found.

21. The CDM system helps me to take command if required by the situation.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4
Hypothesis	Less mental workload would enable the ATCO to spend more capacity on taking initiative if required by the situation. The hypothesis is that there will be no difference in reported mental workload between automation-on and off condition.
Summary of result	No difference expected, no difference found.

22. The CDM system helps me to state my plans when required.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4
Hypothesis	
Summary of result	Not possible to validate with current data set.

23. The CDM system helps me to change plan, with other team members being informed/consulted.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 4
Hypothesis	
Summary of result	Not possible to validate with current data set.

24. With CDM, the task distribution among the team is appropriate.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 5 Ground controller: 5
Hypothesis	Since both agree that distribution was appropriate, it is expected that workload will be lower or equal and Situational Awareness (SA) will be higher or equal in de automation-on condition.
Summary of result	Workload: The departure planner rated for himself a slightly higher workload with automation-on. For the ground controller he rated that workload would be about equal in both conditions. The ground controller himself rated a higher workload for himself and for the departure planner in de automation-off condition. Situational awareness: Differences in SA for both departure planner and ground controller are so small that they are considered equal. Because of the above both are considered validated, but not clearly validated.

25. The CDM system helps me to intervene if task completion deviates from standards.	VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 4 Ground controller: 4
Hypothesis	It is expected that the departure planner will spent less time on radiotelephony in the automation-on condition.
Summary of result	Duration of utterances (Figure 25), utterance duration of departure controller and pilots is shorter in automation-on condition than in automation-off condition Duration of communication (Figure 26), departure clearance takes shorter in automation-on condition than in automation-off condition.

26. The CDM system encourages the team to monitor each other.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 2 Ground controller: 2
Hypothesis	It was expected that in the automation-on condition the departure planner would spend slightly less time monitoring his colleagues.
Summary of result	In the segregated RWYs condition, the departure planner monitors his colleagues as often in the automation-on and -off mode. In the mixed RWYs condition there is no consistent tendency. Not possible to validate with current data set.

27. The CDM system helps me sharing information about developing traffic situations with others.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 4	
Hypothesis	It was expected that in the automation-on condition the departure planner would be slightly more monitoring his colleagues. (So the contrary of the hypothesis of item 26.)	
Summary of result	In the segregated RWYs condition, the departure planner monitors his colleagues as often in the automation-on and -off mode. In the mixed RWYs condition there is no consistent tendency. Not possible to validate with current data set.	

28. The CDM system helps team members alert each other to new events.		NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4	
Hypothesis	It was expected that in both conditions the departure planner would be monitoring his colleagues equally frequently.	
Summary of result	In the segregated RWYs condition, the departure planner monitors his colleagues as often in the automation-on and -off mode. In the mixed RWYs condition there is no consistent tendency. Not possible to validate with current data set.	

29. The CDM system helps the team to synchronise their actions so as not to hinder or delay each other.		VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 5	
Hypothesis	Less hold or stand-by requests in automation-on mode expected for ground controller. For the departure planner no difference is expected.	
Summary of result	Control behaviour (communication) of the departure ATCO (Figure 27, Figure 28 and Figure 29). Asking whether a pilot is ready to start. Requests for start-up which were not immediately cleared (= stand-by) were less frequently observed in automation-on condition than in automation-off condition.	

30. The CDM system makes it easier for the team to identify future overloading.		VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 2 Ground controller: 5	
Hypothesis	Lower workload rating by ground controller and slightly higher workload rating by departure planner in automation-on condition.	
Summary of result	Result was exactly as expected.	

31. It is harder to motivate the team when working with the CDM system.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 3
Hypothesis	
Summary of result	Not possible with current data set. Analyses of speech data would allow a validation of this item.

32. When working with the CDM system, none of the team members needs to take command over the others.	VERIFIED
Q (teamwork questionnaire) rating	Departure planner: 3 Ground controller: 4
Hypothesis	No difference in workload rating.
Summary of result	Only slight (almost equal) differences in reported workload between conditions.

33. With CDM, coordination between team members remains necessary.	NOT VALIDATED
Q (teamwork questionnaire) rating	Departure planner: 1 Ground controller: 4
Hypothesis	Difference expected between conditions automation-on and automation-off with respect to number of monitoring looks and frequency of communication in observation form.
Summary of result	The monitoring look (at ground controller) was observed less frequently in automation-on condition (0.6%) than in automation-off condition (6.7%) for mixed RWYs. However, for segregated RWYs automation-on condition shows more frequent monitoring look (0.5%) than automation-off condition (0.1%). Therefore, with regard to the monitoring look, validated for mixed RWYs, but not for segregated RWYs. Note: Fixation distribution over various AOIs (Figure 12, Figure 13), and monitoring look (Figure 14) have not provided consistent trends.

2. ATCO Ratings on Teamwork Questionnaire

All three ATCOs filled out the teamwork questionnaire after all experiment runs were finished. Their ratings are visualised in Figure 11. Note that the tower controller tends to rate primarily ‘3’. The reason for this is that the tower controller found it not possible to develop a valid opinion about a number of issues. In those cases where he could not make a decision he rated ‘in the middle’.

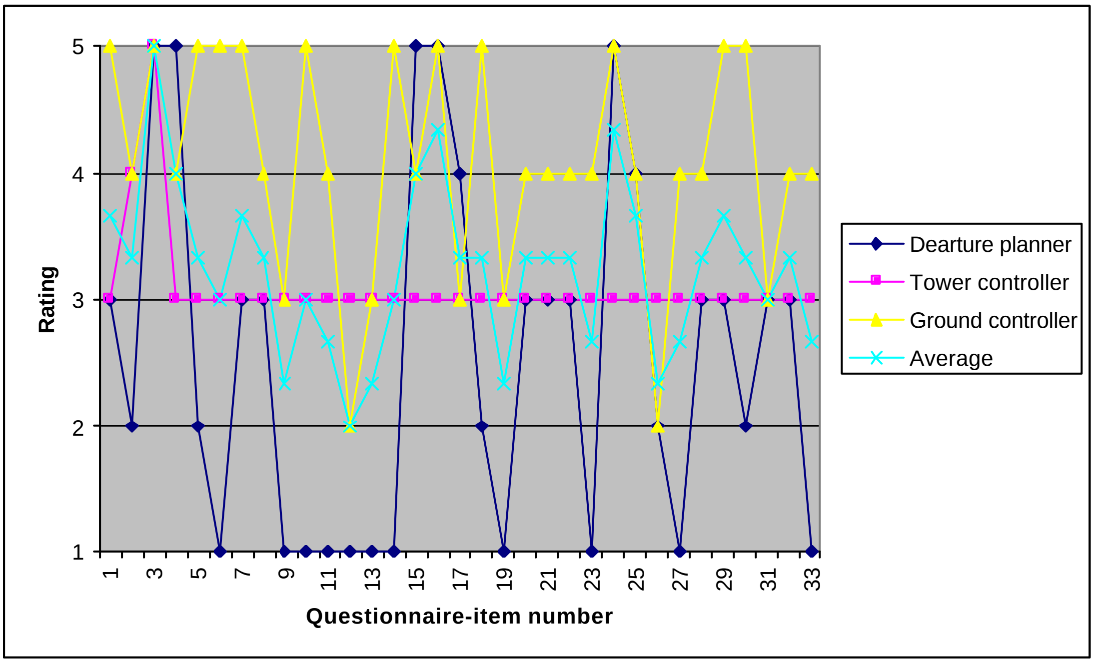


Figure 11: Ratings on teamwork questionnaire (see Appendix B) for all ATCOs

3. Frequencies of Communication as Rated on Observation Form

Table 8: Frequencies of communication as rated on observation form (average over all runs)

Automation-off (average)	**Dep plan**		**Tower**		**Ground**	
	> Tower	>> Ground	< Dep plan	> Ground	< Tower	<< Dep plan
Asks for **information**				1.5	1.7	
Asks someone to perform **action**						
Other requests					1.0	
Gives information/answer	2.0	2.0	2.0	4.3	3.3	2.0
States that s/he will initiate **action**				2.0		
Suggests action or **solution**				1.0	1.0	
Other announcements				1.0	2.0	
Acknowledgements				1.0	2.0	

Automation-on (average)	**Dep plan**		**Tower**		**Ground**	
	> Tower	>> Ground	< Dep plan	> Ground	< Tower	<< Dep plan
Asks for **information**			1.0	1.0	2.0	
Asks someone to perform **action**					1.0	
Other requests						
Gives information/answer				1.0	1.5	
States that s/he will initiate **action**						
Suggests action or **solution**						
Other announcements				1.0	1.0	
Acknowledgements			2.0	2.0	1.0	

<u>Note</u>: In both conditions there is, over all, very little communication taking place.

4. Situation Awareness Rating

Table 9: Situation awareness rating

	Run 1			**Run 2**			**Run 3**		
Rated by:	Dep plan	Tower	Ground	Dep plan	Tower	Ground	Dep plan	Tower	Ground
Value for:									
Tower	8	8	7	8	7	8	10	8	7
Dep Plan	8		7	10		7	10		8
Ground	4	8	6	5	8	8	10	8	8
	Run 4			**Run 5**			**Run 6**		
Rated by:	Dep plan	Tower	Ground	Dep plan	Tower	Ground	Dep plan	Tower	Ground
Value for:									
Tower	9	8	9	9	8	8	9	8	8
Dep Plan	9		9	9		8	9		8
Ground	9	8	9	9	8	8	9	8	7

5. Mental Workload Rating

Table 10: Mental workload rating

	Run 1			**Run 2**			**Run 3**		
Rated by:	Dep plan	Tower	Ground	Dep plan	Tower	Ground	Dep plan	Tower	Ground
Value for:									
Tower	4.3	5.8	7	5.5	7.1	4	5.8	5.8	6
Dep Plan	4.6		4	5.5		4	1.9		7
Ground	10.1	8.6	10	8.4	5.8	6	6.6	4	7
	Run 4			**Run 5**			**Run 6**		
Rated by:	Dep plan	Tower	Ground	Dep plan	Tower	Ground	Dep plan	Tower	Ground
Value for:									
Tower	4	3.8	4	2.7	7.1	7	3.5	5.8	7
Dep Plan	2.5		4	5		7	5		7
Ground	7.1	3.8	4	8	7.1	9	8.1	7.1	9

6. SASHA Questionnaire

Q1. Did you have the feeling that you were ahead of the traffic, able to predict the evolution of the traffic?
Q2. Did you have the feeling that you were able to plan and organise your work as you wanted?
Q3. Have you been surprised by an aircraft call that you were not expecting?
Q4. Did you have the feeling of starting to focus too much on a single problem?
Q5. Did you forget to act on any aircraft?
Q6. Did you have any difficulty finding an item of (static) information?
Q7. Do you think that the system provided you with useful information?
Q8. Were you paying too much attention to the functioning of the system?
Q9. Did the system help you to have a better understanding of the situation?
Q10. Finally, how would you rate your overall situation awareness during the exercise?

1 = never, 5 = always or 1 = poor, 5 = very good

Table 11: Results per run and controller

Run 1			**Run 2**			**Run 3**		
Dep plan	Tower	Ground	Dep plan	Tower	Ground	Dep plan	Tower	Ground
4	5	3	5	4	4	5	5	4
4	5	3	5	3	4	5	5	5
1	2	2	1	2	1	1	2	1
1	2	2	1	2	1	1		2
1	2	2	1	2	1	1	2	2
1	2	1	1	2	1	1	2	1
5	4	1	5	4	4	3	4	5
2	3	2	1	3	2	1	2	1
5	4	4	5	3	5	3	4	4
5	4	2	5	4	3	5	4	4
Run 4			**Run 5**			**Run 6**		
Dep plan	Tower	Ground	Dep plan	Tower	Ground	Dep plan	Tower	Ground
5	5	5	3	5	4	4	5	4
5	5	5	3	5	4	5	4	4
1	2	1	2	1	2	4	2	2
1	1	1	1	1	1	1	3	1
1	1	1	1	1	1	1	2	2
1	1	1	1	1	1	1	2	1
5	3	5	3	5	5	3	5	5
1	2	1	1	2	2	1	2	1
5	4	4	2	5	4	4	5	4
5	4	4	2	4	4	4	4	3

7. Fixation Distribution among AOIs

Figure 12 shows the fixation distribution among AOIs (based on EPOG data). In all runs the left and middle monitors were looked for more than 80 % of the whole time.

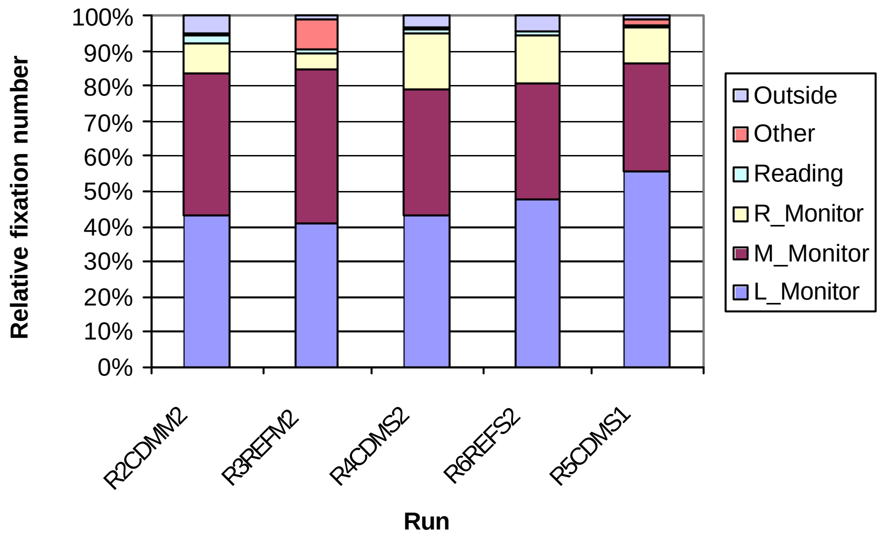


Figure 12: Fixation distribution on various AOIs (based on fixation number in EPOG)

The AOI data evaluated manually from the eye tracking data shows partly different tendency (see Figure 13 below).

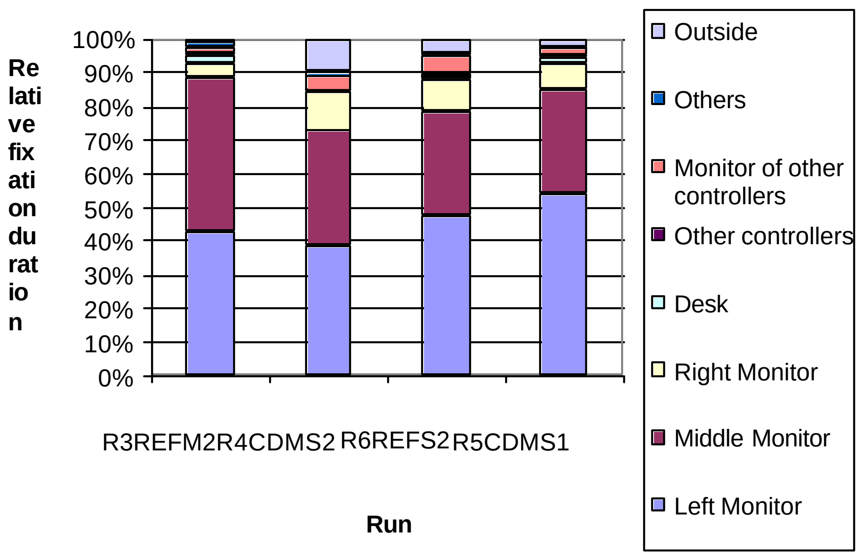


Figure 13: Fixation distribution on various AOIs (based on video analysis)

Monitoring look is divided furthermore into two groups, namely monitoring of other controllers and that of the monitors of other controllers. The monitoring of other controllers was observed very rarely. The departure controller reported after the experiment that he would more frequently have communicated with the ground controller, if the ground controller were next to him. The comparison between CDM on and CDM off shows that the monitoring look on the other controllers is more frequently in CDM-on conditions than in CDM-off conditions.

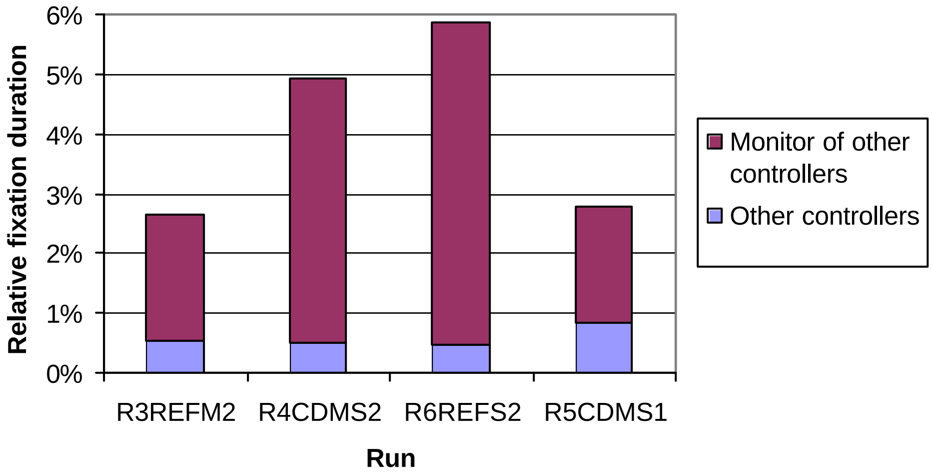


Figure 14: Monitoring look (based on fixation duration)

Table 12: EPOG for two complete (comparable) runs (segregated runway use)

Run 4 (automation on, scenario 2, segregated runway use)	**Run 6** (automation off, scenario 2, segregated runway use)
Blink rate (blink/min) 48.41	Blink rate (blink/min) 24.5
Fixations on arrivals table 2.15 %	Fixations on arrivals table 7.16 %
Fixations on departure table 15.15 %	Fixations on departure table 15.24 %
Fixations on EFS top 3.56 %	Fixations on EFS top 0.98 %
Fixations on EFS middle 2.61 %	Fixations on EFS middle 27.59 %
Fixations on EFS bottom 0.05 %	Fixations on EFS bottom 0.61 %
Fixations on EFS right 4.49 %	Fixations on EFS right 8.91 %
Fixations on airport map 5.5 %	Fixations on airport map 8.48 %
Fixations on outside (projection screen) 17.04 %	Fixations on outside (projection screen) 10.01 %
Fixations on other predefined areas 9.31 %	Fixations on other predefined areas 2.62 %

Table 12 compares the automation-on versus automation-off situation for simulated runs both with segregated runway use. The blink rate (the average amount of eye blinks made by the departure planner per minute) is indicated. The strong correlation between blink rate en (visual) workload makes blink rate an interesting indicator. (The lower the blink rate the higher the visual workload (see Harris *et al.*, 1986; Stern *et al*., 1984; Stern & Kelly, 1984; Stern, 1994; Goldstein *et al.*, 1985; Wilson, 1987, 1993). Further the percentage of fixations on particular areas of interest is given. Notice that when automation is off the departure planner makes more fixations on the EFS monitor compared to the situation when automation was switched on. In the automation-on situation the departure planner spent more time looking outside (at the projection screen). The data loss in run 4 was higher than in run 6.

Table 13. EPOG for two complete (comparable) runs (mixed runway use)

Run 2 (automation on, scenario 2, mixed runway use)	**Run 3** (automation off, scenario 2, mixed runway use)
Blink rate (blink/min) 41.10	Blink rate (blink/min) 8.64
Fixations on arrival-departure table 10.49 %	Fixations on arrivals table 22.04 %
Fixations on EFS top 1.91 %	Fixations on EFS top 0.29 %
Fixations on EFS middle 1.64 %	Fixations on EFS middle 8.63 %
Fixations on EFS bottom 0.02 %	Fixations on EFS bottom 0.61 %
Fixations on EFS right 2.86 %	Fixations on EFS right 5.66 %
Fixations on airport map 1.66 %	Fixations on airport map 1.75 %
Fixations on outside (projection screen) 14.07 %	Fixations on outside (projection screen) 1.3 %
Fixations on other predefined areas 10.96 %	Fixations on other predefined areas 8.75 %

Table 13 compares the automation-on versus off situation for simulated runs both with mixed runway use. The blink rate is indicated. Further the percentage of fixations on particular areas of interest is given. Notice that when automation is off the departure planner makes more fixations on the EFS monitor compared to the situation when automation was switched on. In the automation on situation the departure planner spent more time looking outside (at the projection screen). The data loss was higher than in Table 12 but for both runs about the same.

Table 12 and Table 13 both indicate that when automation is on there is the tendency to look less often at arrivals and/or departures and less often at the EFSs as well. This leaves more time for looking outside (in the current low workload scenarios). Since it is statistically speaking difficult to project these results based on just one ATCO, we cannot explain the implications of this ‘trend’ on real-life situations.

8. Change in the Fixation Distribution in the Course of Time

The change in the fixation distribution in the course of time is an interesting point. It could show certain tendency. As Figure 15 to Figure 18 show, however, there is no consistent tendency in the time line. It rather depends on the situation.

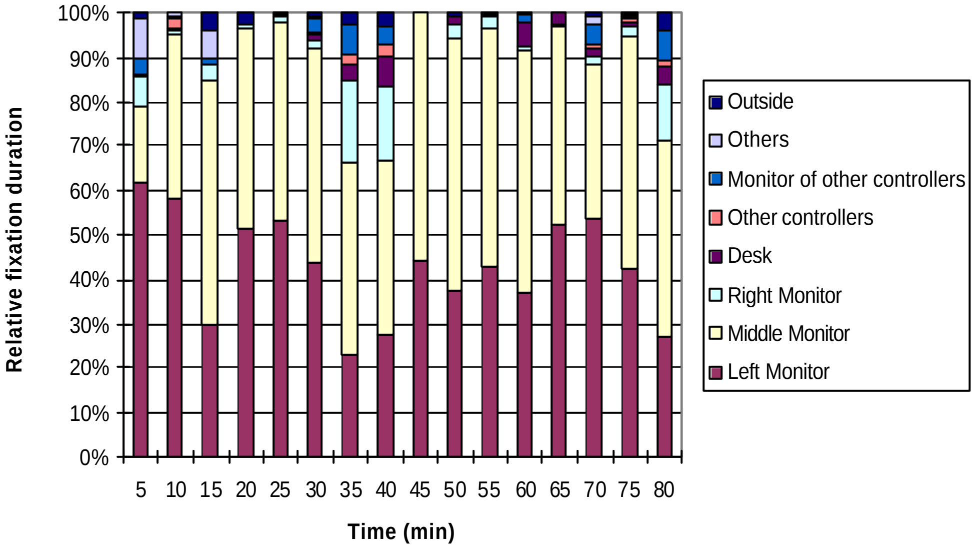


Figure 15: Change in the fixation distribution (R3REFM2)

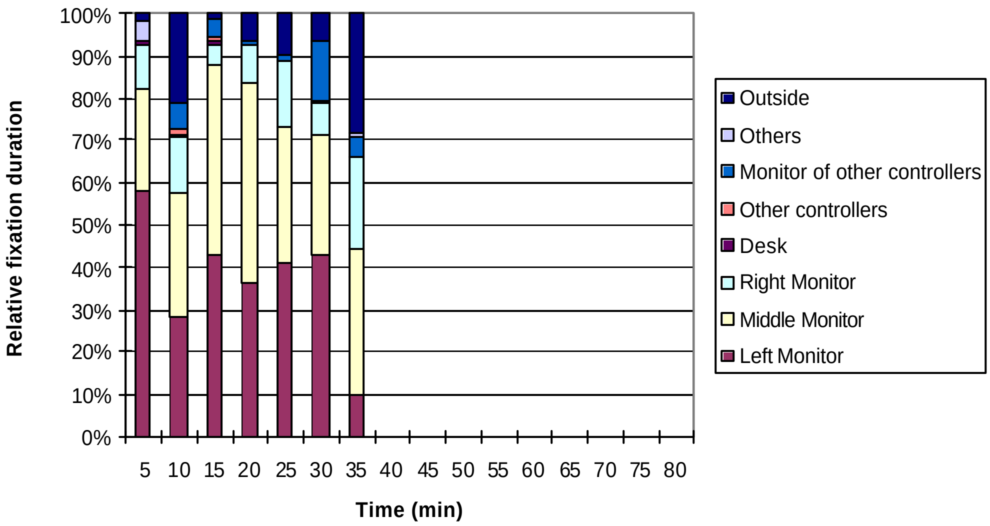


Figure 16: Change in the fixation distribution (R4CDMS2)

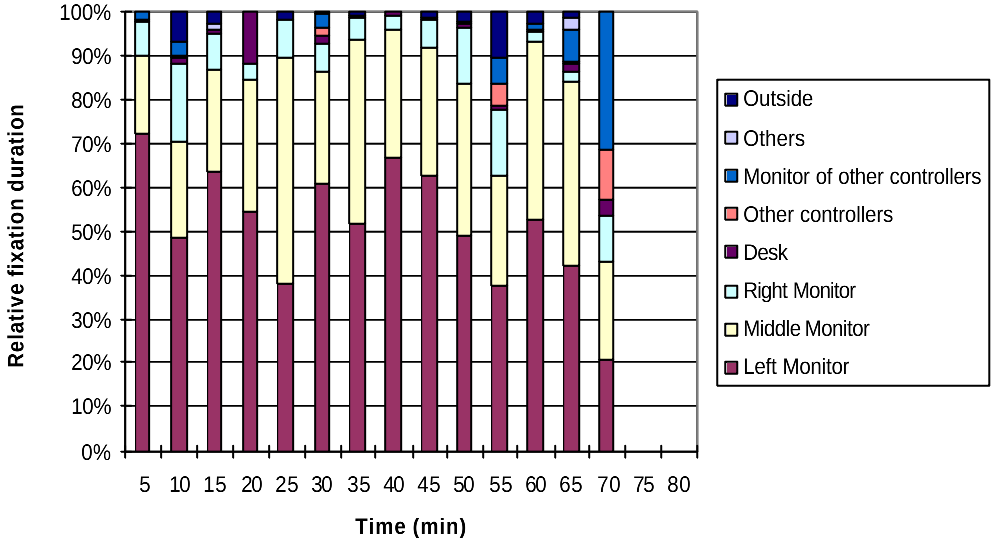


Figure 17: Change in the fixation distribution (R5CDMS1)

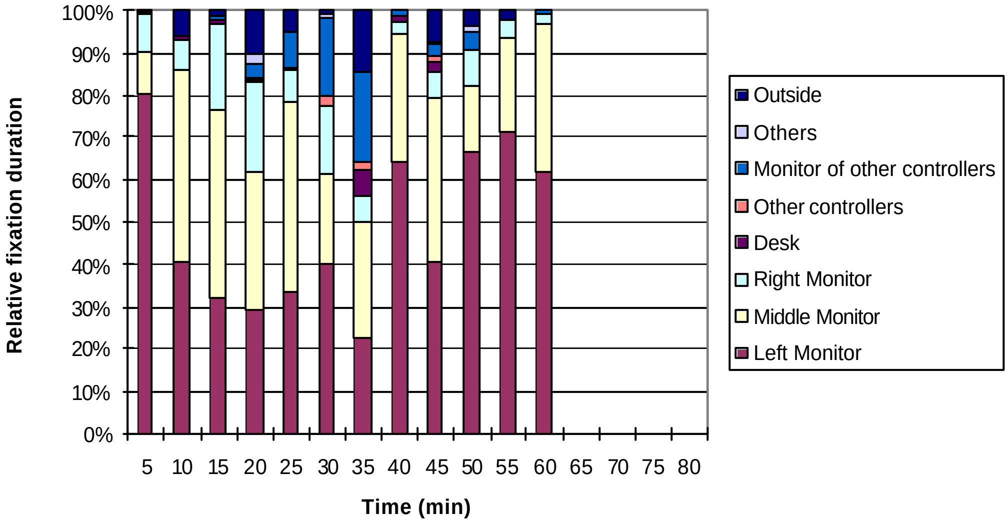


Figure 18: Change in the fixation distribution (R6REFS2)

9. Fixation Frequency within each AOI

Fixation frequency or dwell frequency in certain time span and fixation frequency per dwell are indices for the stability of the eye movement.

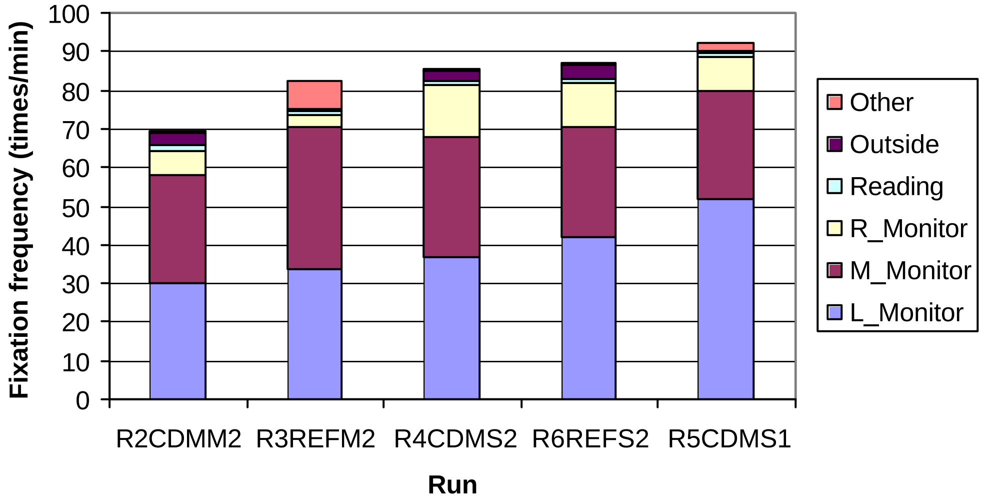


Figure 19: Fixation frequency (per min) (EPOG)

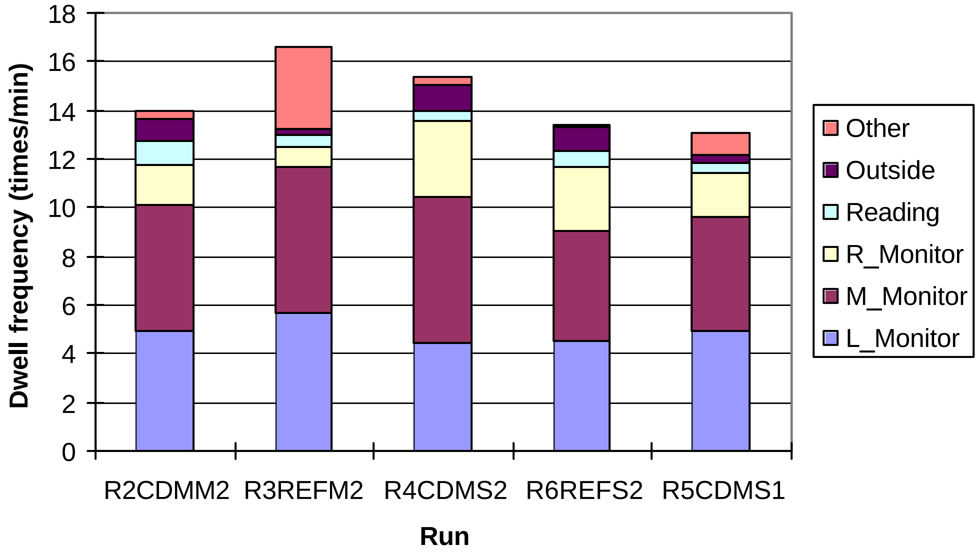


Figure 20: Dwell frequency (per min) (EPOG)

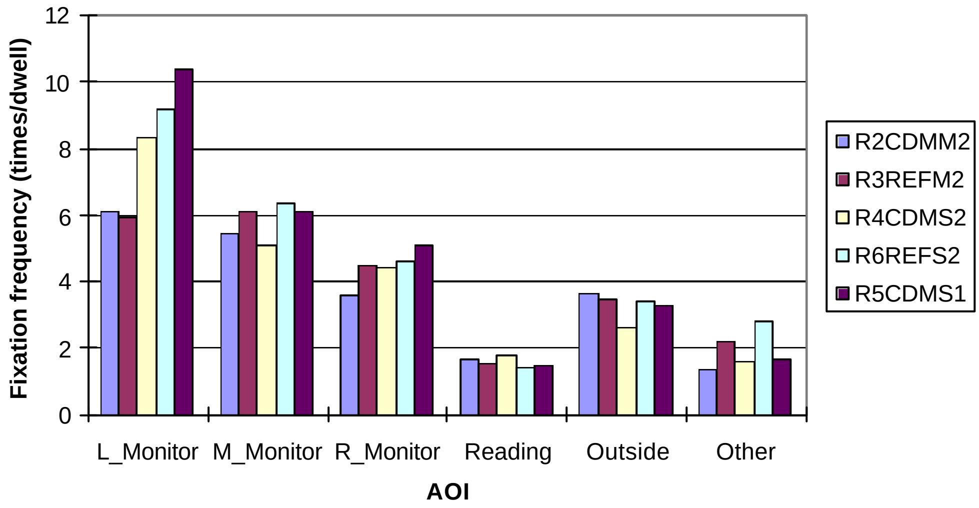


Figure 21: Fixation frequency (per dwell) (EPOG)

10. Fixation Transition between Different Monitors

As Figure 22 shows, fixation transition between the middle monitor and the right monitor was frequently observed in CDM-on condition than in CDM-off condition. With regard to other combination, there was no salient difference between CDM-on condition and CDM-off condition.

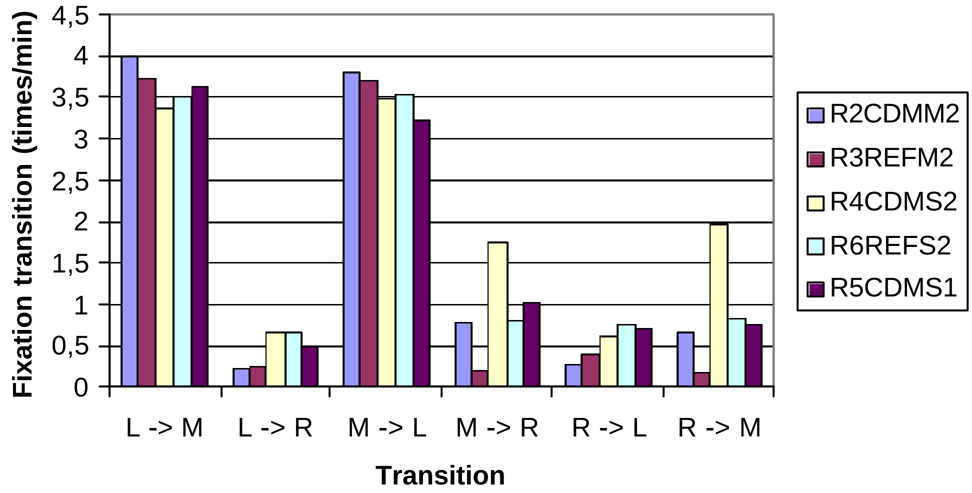


Figure 22: Fixation transition between different monitors (EPOG)

11. Entropy in Eye Movement

Entropy shows whether there are some patterns in the eye movement. It is possible that the departure controller looks at the monitors or other things in a consistent way.

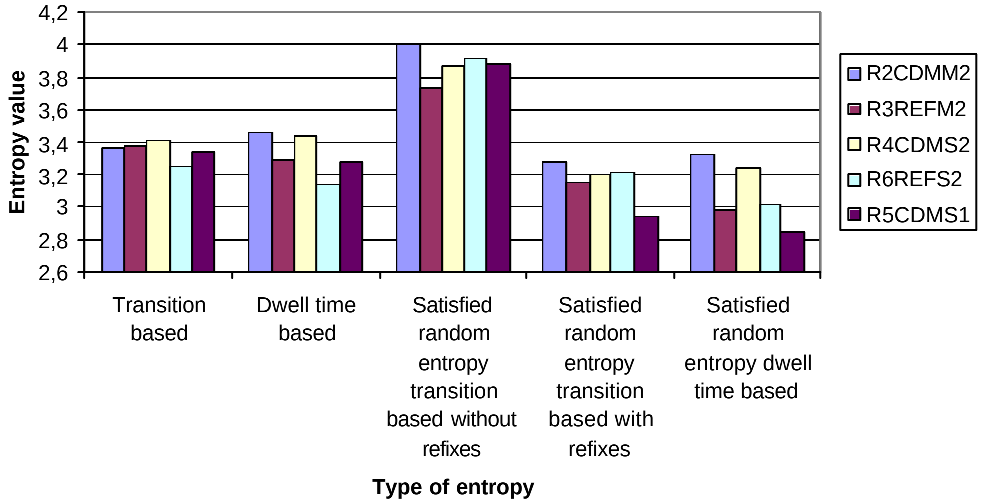


Figure 23: Entropy (EPOG)

12. Reaction Time to the Information Given by a Pilot

If a pilot started to talk to the departure controller, the departure controller turned immediately to the left monitor that displays the EFSs.

Figure 24 shows the reaction time of the departure controller to the utterance of pilots. The reaction by eye movement to the left monitor was quicker in the runs with CDM on than in those without CDM functionality.

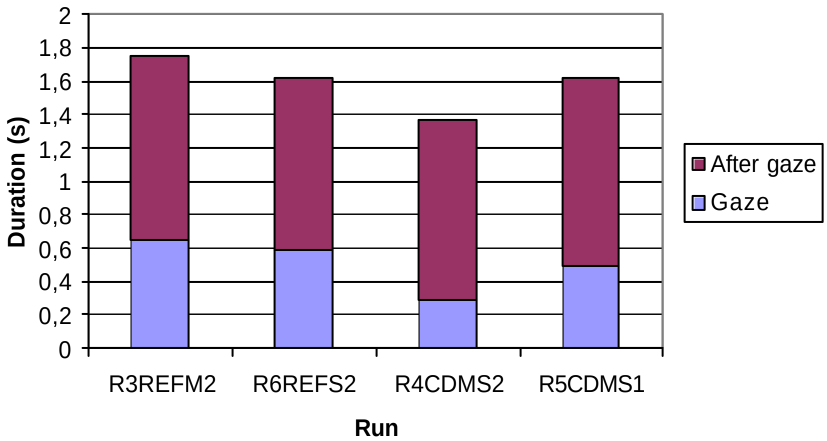


Figure 24: Reaction time of the departure planner to the utterance of pilots

13. Duration of Utterance

Such highly standardized utterances produce not so much differences in the duration. However, as Figure 25 shows, the duration of each utterance is shorter with CDM than that without CDM. Without CDM the departure controller murmured sometimes and it causes longer duration of utterances. The longer duration of utterances by pilots is assumingly caused by the thinking process for the reaction.

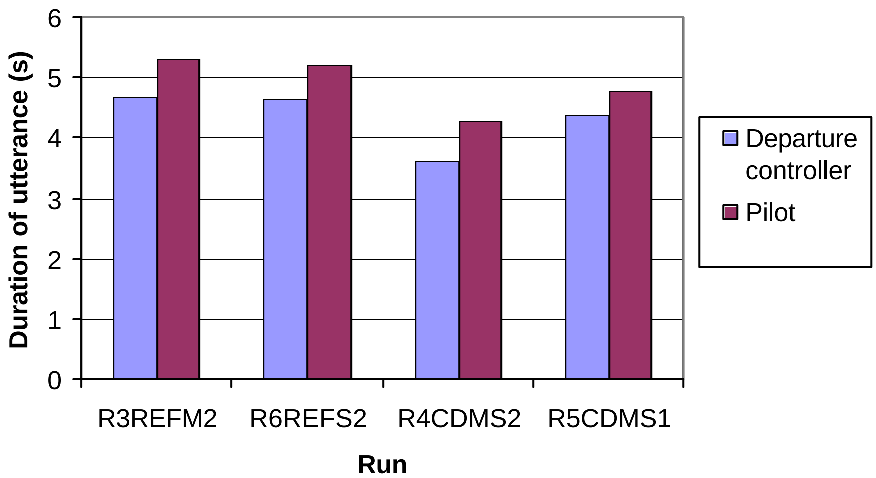


Figure 25: Duration of the utterances of departure controller and pilot

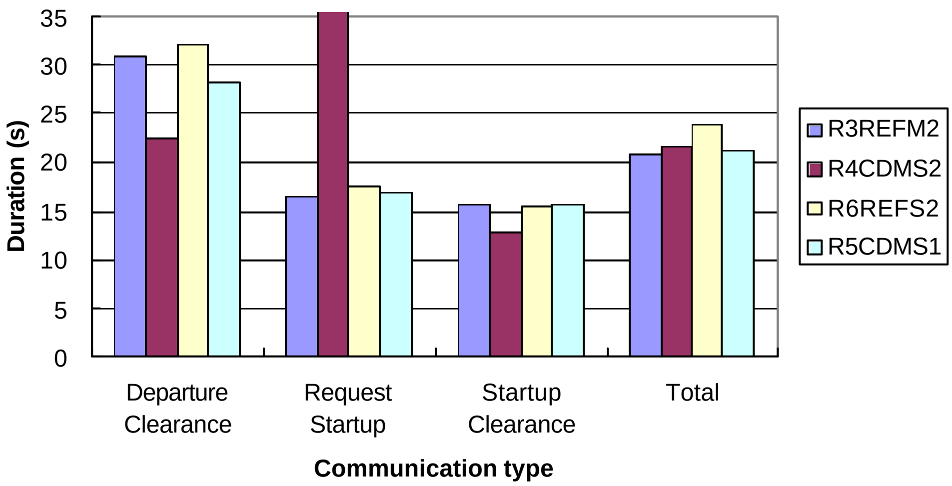


Figure 26: Duration of communication between departure controller and pilot

14. Communication Behaviour

Which type of communication is how often required is fully depending on the situation. Figure 27 shows the observed frequency of each communication type for the whole simulation process, while Figure 28 shows such frequency per minute. The type 'request start-up / ready to start?' was observed less/seldom in CDM-on condition than in CDM-off condition.

With regard to this communication type, the order for 'stand-by' by departure controller was given seven times in R3REFM2, one time in R4CDMS2, and eight times in R5CDMS1 and R6REFS2.

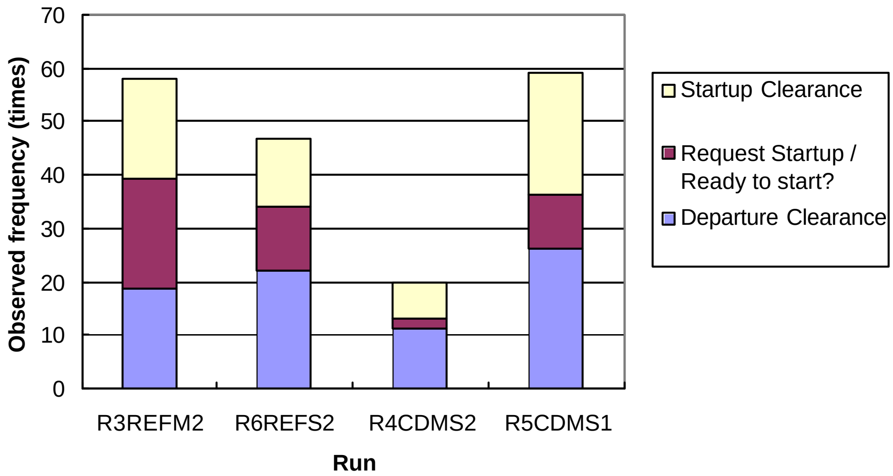


Figure 27: Observed communication types (for whole simulation process)

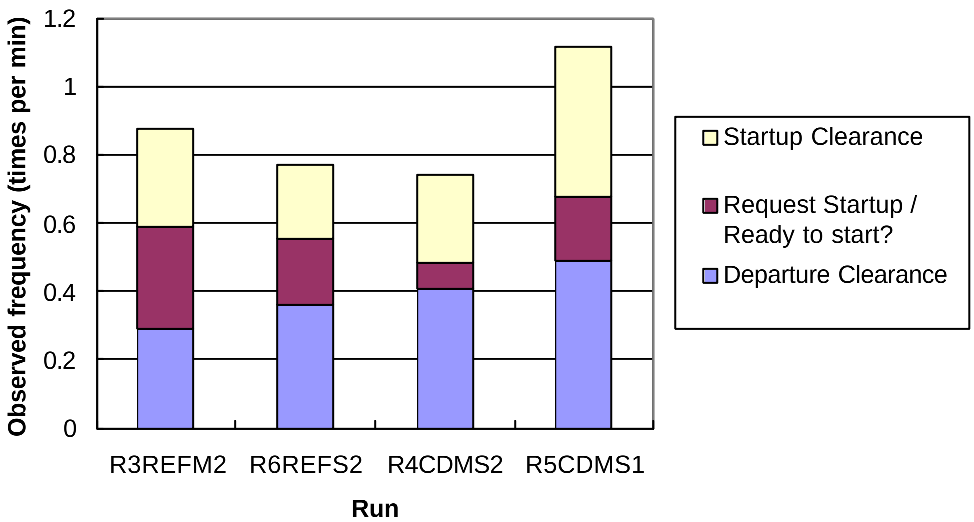


Figure 28: Observed communication types (per min)

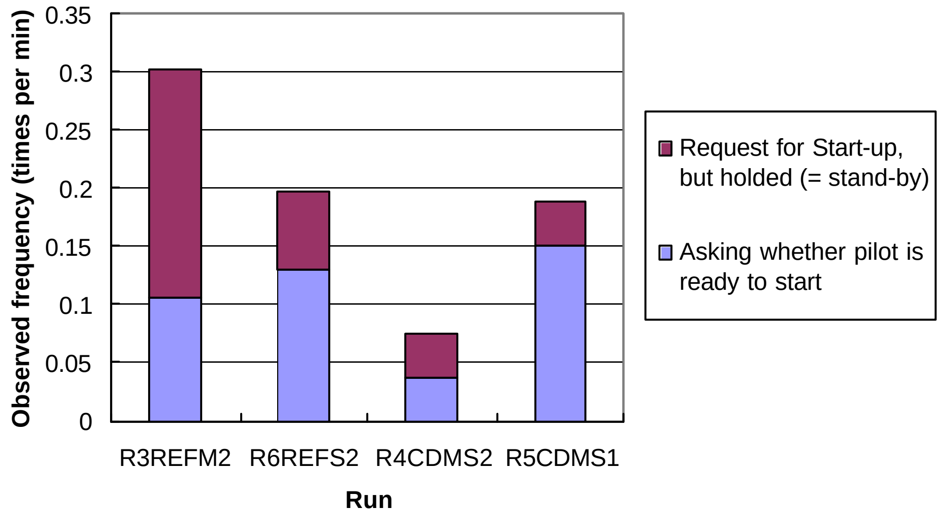


Figure 29: Observed communication types (per min) (2)