EUROPEAN ORGANISATION
FOR THE SAFETY OF AIR NAVIGATION

EUROCONTROL

This Document is issued as EATMP Method and Tool. The contents are not mandatory. They provide information and explanation or may indicate best practice.

The Development of Situation Awareness Measures in ATM Systems

Edition Number	:	1.0
Edition Date	:	27.06.2003
Status	:	Released Issue
Intended for	:	EATMP Stakeholders

EUROPEAN AIR TRAFFIC MANAGEMENT PROGRAMME

DOCUMENT CHARACTERISTICS

TITLE

The Development of Situation Awareness Measures in ATM Systems

		EATMP Infocentre Reference:	030317-04
Document Identifier		**Edition Number:**	1.0
HRS/HSP-005-REP-01		**Edition Date:**	27.06.2003

Abstract

This report is concerned with the measurement of controller's Situation Awareness (SA). Two measures of SA are proposed: the first, known as 'SA for SHAPE on-Line (SASHA_L)' is a query technique based on an existing measure; the second, known as 'SA for SHAPE Questionnaire (SASHA_Q)', is a questionnaire technique using carefully chosen questions that focus on key elements of SA which controllers have identified themselves. A wide-ranging review of nine current SA measures is also included.

The report has been prepared as part of the 'Solutions for Human-Automation Partnerships in European ATM (SHAPE)' Project being carried out by the ATM Human Resources Unit of EUROCONTROL, later renamed 'Human Factors and Manpower Unit (DIS/HUM)', and today known as 'Human Factors Management Business Division (DAS/HUM)'.

Keywords

Air Traffic Management (ATM) System	Automation
Human Factors	Measure
Measurement	Real-time Simulation
SA for SHAPE (SASHA)	Situation Awareness (SA)
Solutions for Human-Automation Partnerships in European ATM (SHAPE)	Validation

Contact Persons	Tel	Division
Oliver STRAETER, SHAPE Project Leader	+32 2 7295054	Human Factors Management Business Division (DAS/HUM)
Michiel WOLDRING, Manager, HRS Human Factors Sub-Programme (HSP)	+32 2 7293566	Human Factors Management Business Division (DAS/HUM)

Authors

E. Jeannot, C. Kelly and D. Thompson

STATUS, AUDIENCE AND ACCESSIBILITY

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

ELECTRONIC SOURCE

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

EATMP 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: eatmp.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
Please make sure that the EATMP Infocentre Reference is present on page ii.		
SHAPE Project Leader	O. STRAETER	12.06.03
Chairman HRT Human Factors Focus Group (HFFG)	V. S. M. WOLDRING	12.06.2003
Manager EATM Human Resources Programme (HRS-PM)	M. BARBARINO	12.06.2003
Chairman EATM Human Resources Team (HRT)	A. SKONIEZKI	12.06.2003
Senior Director EATM Service Business Unit (SD)	W. PHILIPP	18.06.03

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	17.09.2001		Working Draft	All
0.2	07.11.2001		Draft	All
0.3	13.02.2002		Approval by HFSG7	All
0.4	30.08.2002		Proposed Issue for HRT18	All (document configuration)
1.0	27.06.2003	030317-04	Released Issue (agreed for release at HRT18 on 9-10 Oct 2002)	All (final editorial changes & document configuration)

CONTENTS

EXECUTIVE SUMMARY

This report describes the development of two human factors tools for measuring Situation Awareness (SA) in Air Traffic Management (ATM) systems. The measures are primarily concerned with controllers' SA when using computer-assistance 'tools' and other forms of automation support, which are expected to be major components of future ATM systems.

The report has been prepared as part of the 'Solutions for Human-Automation Partnerships in European ATM (SHAPE)' Project being carried out by the ATM Human Resources Unit of EUROCONTROL, later renamed 'Human Factors and Manpower Unit (DIS/HUM)' and today known as 'Human Factors Management Business Division (DAS/HUM)'.

Three other documents provide human factors guidelines for facilitating and fostering human trust in ATM systems; one proposes a literature review, the other techniques for measuring trust and the third one gives trust principles (see EATMP, 2003a, 2003b and 2003c). A further report deals with the teamwork and automation (currently under preparation). Four additional human factors issues are also in the SHAPE overall objectives: recovery from system failure, workload and automation, future controller skill-set requirements, and experience and age (see EATMP, 2003d, in press).

Section 1, 'Introduction', outlines the background to the project, and the objectives and scope of the report.

Section 2, 'Situation Awareness Background', defines what is meant by SA and describes some theoretical approaches.

Section 3, 'Situation Awareness and ATM Systems', reviews research into SA that has been carried out specifically in the ATM domain. The topic of team SA is briefly reviewed. The meaning and interpretation of SA from the controllers' point of view is also discussed.

Section 4, 'Situation Awareness Measures', provides a detailed review of nine measures of SA. For each measure, the background theory, method, and the scoring of data are presented. The advantages and disadvantages of each measure are also summarised. Finally, the implications for SHAPE are discussed and two SA measures are proposed.

Section 5, 'Situation Awareness Measures for SHAPE', describes two measures of SA that are proposed for SHAPE. The first measure, 'SA for SHAPE on-Line' (SASHA_L), is a based on the development of an existing technique. The second measure, 'SA SHAPE questionnaire' (SASHA_Q), is a questionnaire-based measure. The validity of the proposed measures is discussed.

Section 6, 'Conclusions', presents the conclusions of the report. Amongst the conclusions drawn is the importance of involving Subject Matter Experts (SMEs) for implementing the on-line SA measure. In addition, the importance of proper training is emphasised.

Section 7, 'Recommendations', presents the recommendations of the report. Amongst the recommendations made is that further work is necessary to validate the two proposed measures of SA.

A Bibliography, Further Reading, a list of the Abbreviations and Acronyms used in this report, and Acknowledgements are also provided.

Finally, Appendix A provides the SASHA Questionnaire while Appendix B gives the instructions for SASHA on-Line.

1. INTRODUCTION

1.1 Purpose

The purpose of this report is to provide a human factors measurement technique for measuring Situation Awareness (SA) in ATM systems. The measure is primarily concerned with controllers' SA when using computer-assistance 'tools' and other forms of automation support, which are expected to be major components of future ATM systems.

1.2 Scope

The concept of SA is not new and much research has been carried out on the subject since it was first conceived in the late eighties, in the context of enhancing pilot performance in fighter aircraft. A variety of techniques for measuring SA have also been developed over the past ten years. Therefore, at the outset, it is assumed that this substantial body of existing research can be built upon for the purposes of SHAPE without the need to develop a new measure from scratch. Similarly, although relevant literature is cited and discussed, it is beyond the scope of this report to carry out an extensive review. For detailed reviews the reader is referred to Dominguez *et al.* (1994), and Endsley and Garland (2000).

The SA measure is intended principally for use in real-time simulations conducted at the EUROCONTROL Experimental Centre (EEC), Brétigny, France, particularly those concerned with investigating the introduction of ATM Decision Support Tools (DST). Hence, this report is aimed at project leaders, project analysts and other project staff who are concerned with the human factors aspects of such simulations.

1.3 Background

The work on situational awareness in this module is embedded in a larger project called 'Solutions for Human-Automation Partnerships in European ATM (SHAPE)'. The SHAPE Project started in 2000 within the Human Factors Sub-Programme (HSP) of the EATMP Human Resources Programme (HRS) conducted by the ATM Human Resources Unit of EUROCONTROL, later renamed 'Human Factors and Manpower Unit (DIS/HUM)' and today known as 'Human Factors Management Business Division (DAS/HUM)' (see EATMP, 2000).

SHAPE is dealing with a range of issues raised by the increasing automation in European ATM. Automation can bring success or failure, depending on whether it suits the controller. Experience in the introduction of automation into cockpits has shown that, if human factors are not properly considered, 'automation-assisted accidents' may be the end result.

Seven main interacting factors have been identified in SHAPE that need to be addressed in order to ensure harmonisation between automated support and the controller:

- Trust: The use of automated tools will depend on the controllers' trust. Trust is a result of many factors such as reliability of the system and transparency of the functions. Neither mistrust nor complacency are desirable. Within SHAPE guidelines were developed to maintain a correctly calibrated level of trust (see EATMP, 2003a, 2003b, 2003c).

- Situation Awareness (SA): Automation is likely to have an impact on controllers SA. SHAPE developed a method to measure SA in order to ensure that new systems do not distract controllers' situation awareness of traffic too much (covered by this document).

- Teams: Team tasks and performance will change when automated technologies are introduced (team structure and composition change, team roles are redefined, interaction and communication patterns are altered). SHAPE has developed a tool to investigate the impact of automation on the overall team performance with a new system (currently under preparation).

- Skill set requirements: Automation can lead to both skill degradation and the need for new skills. SHAPE identifies new training needs, obsolete skills, and potential for skill degradation aiming at successful transition training and design support (currently under preparation).

- Recovery from system failure: There is a need to consider how the controller will ensure safe recovery should system failures occur within an automated system (currently under preparation).

- Workload: With automation human performance shifts from a physical activity to a more cognitive and perceptual activity. SHAPE is developing a measure for mental workload, in order to define whether the induced workload exceeds the overall level of workload a controller can deal with effectively (currently under preparation).

- Ageing: The age of controllers is likely to be a factor affecting the successful implementation of automation. Within SHAPE this particular factor of human performance, and its influence on controllers' performance, are investigated. The purpose of such an investigation is to use the results of it as the basis for the development of tools and guidance for supporting older controllers in successfully doing their job in new automated systems (see EATMP, 2003d, in press). Note that an additional report providing a questionnaire-survey throughout the Member States of EUROCONTROL is currently under preparation.

These measures and methods of SHAPE support the design of new automated systems in ATM and the definition of training needs. It also facilitates the preparation of experimental settings regarding important aspects of human performance such as potential for error recoveries or impacts of human performance on the ATM capacity.

The methods and tools developed in SHAPE will be complied in a framework in order to ease the use of this toolkit in either assessing or evaluating the impact of new systems on the controller performance, efficiency and safety. This framework will be realised as a computerised toolkit and is planned to be available end of 2003.

1.4 Structure

The document is divided into seven sections. In Section 2, following this 'Introduction', background information on SA is discussed. A number of approaches to model SA are discussed. In Section 3 the treatment of SA within the context of ATC and ATM systems is reviewed. The controllers' view of SA is considered, including information gathered from interviewing operational controllers. Research on the topic of team SA is also reviewed. In Section 4, various techniques for the measurement of SA are reviewed; the implications for SHAPE are presented. Based on this information two SA measures for SHAPE are proposed in Section 5. The conclusions of the report are given in Section 6 and recommendations in Section 7.

Page intentionally left blank

2. SITUATION AWARENESS BACKGROUND

2.1 What is Situation Awareness?

Put simply, situation awareness (SA) means knowing what is going on around you. More specifically, in the context of complex operational environments SA is concerned with the person's knowledge of particular task-related events and phenomena. For example, for a fighter pilot SA means knowing about the threats and intentions of enemy forces as well as the status of his/her own aircraft. For an air traffic controller, SA means (at least partly) knowing about current aircraft positions and flight plans and predicting future states so as to detect possible conflicts. Therefore, in operational terms, SA means having an understanding of the current state and dynamics of a system and being able to anticipate future change and developments.

A general definition of SA is that it is *the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future* (Endsley, 1988). This basic definition has been extended by Dominguez *et al.* (1994), who state that SA needs to include the following four specific pieces of information:

- extracting information from the environment;
- integrating this information with relevant internal knowledge to create a mental picture of the current situation;
- using this picture to direct further perceptual exploration in a continual perceptual cycle; and
- anticipating future events.

Taking these four elements into account, SA is defined as *the continuous extraction of environmental information, the integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events* (Dominguez *et al.*, 1994).

Situation awareness is the continuous extraction of environmental information, the integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events.

Three additional remarks need to be made concerning the definition of SA:

- First, the reference to 'mental picture' in the above definition highlights a point that SA is sometimes seen as merely a different label for what is referred to in the ATM domain as the controller 'having the picture'. This is an important point that will be returned to later in the report when considering controllers' views of SA.
- Second, as a corollary of the last point, SA is also viewed as a form of 'mental model'. This is further discussed in 2.2.
- Third, because of its origins in the aviation domain, and particularly the focus on pilots, SA has primarily been considered in relation to the individual operator or controller. However, in recent years research into **team** SA has emerged, which is of direct relevance to ATC (also see 3.2).

2.2 Theories of Situation Awareness

Despite the abundance of research literature about SA, surprisingly few theories and models of SA have been advanced. Indeed, in a recent review of SA from the perspective of cognitive psychology Durso and Gronlund (2000) concluded *this young field was able to supply fewer empirical findings than is needed to develop a coherent theory* (p. 285).

That being said, the most well-known theory, or certainly the most widely publicised, is that of Endsley (1988, 1995a, 1995b, 2000). She proposed a framework model of SA based on information-processing theory. The Endsley Model, as shown in Figure 1, depicts SA as an internal model derived from the environment that is separate from, and precedes **decision-making** and performance. As can also be seen, a range of individual and task/system factors (e.g. memory, goals, workload and automation) affects SA.

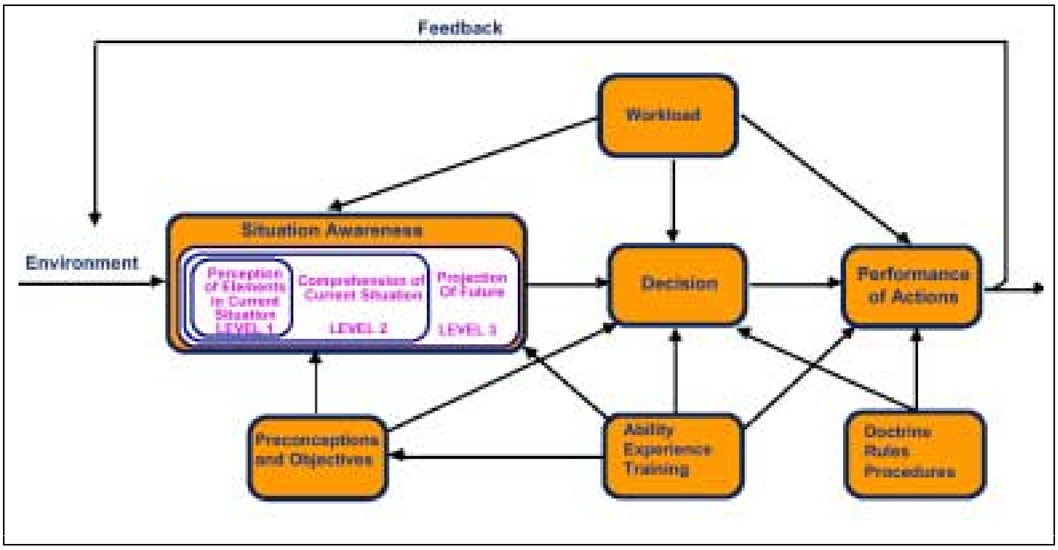

Figure 1: Model of SA in human decision-making (Endsley & Rodgers, 1994)

Another important aspect of Endsley's Model is that SA can be split into three levels of information processing. These levels, as shown in Figure 1, are:

1. **Level 1 SA - Perception.** This is the first fundamental step in SA and involves perceiving and attending to important cues or 'elements' in the environment.

2. **Level 2 SA - Comprehension**. This step goes beyond mere perception and involves integrating different pieces of (Level 1) data and information and determining their operational relevance.

3. **Level 3 SA - Projection**. This highest level of SA involves being able to anticipate future events and their implications based on the (Level 2) understanding of the environment. Level 3 SA allows for timely **decision-making.**

This three-level model of SA underlies the well-known 'Situation Awareness Global Assessment Technique (SAGAT)' measure of SA developed by Endsley. This measure is discussed in 4.2.

As noted in 2.1, SA has been treated by some researchers as a form of mental model. The existence of mental models has long been taken for granted in the research of human-machine systems (e.g. Edwards & Lees, 1974; Gentner & Stevens, 1983) and human-computer interaction (Carroll & Olson, 1987). Indeed, there is a considerable body of research literature on the subject, which is beyond the scope of this report to review. However, a few brief points can be made.

A mental model is most commonly defined as a representation, in a person's head, of a physical system and/or software. For example, according to Carroll and Olson (*op. cit.*), a mental model is *a rich and elaborate structure, reflecting the user's understanding of what the system contains, how it works, and why it works that way. It can be conceived as knowledge about the system sufficient to permit the user to mentally try out actions before choosing one to execute* (p. 12). In this sense, a mental model is the underlying knowledge on which SA is based (Mogford, 1997). Similarly, Endsley (2000) states that *SA, or the situation model, is the current state of the mental model* (p. 16).

Other authors have expressed disagreement with this notion of SA as a mental model arguing that the environment is too dynamic and that SA is a result of **interaction** with the environment (Smith & Hancock, 1995). In fact, the authors define SA as *adaptive, externally directed consciousness* (p. 138). However, what is clear is that the two concepts - mental model and SA - are closely related. Moreover, that relationship also explains why the similarity or lack of it, between SA and the controller's 'picture' (see 3.3) is a subject of continuing debate.

In contrast to Endsley's information processing approach, other researchers have emphasised the importance of perception. Drawing on the work of Neisser (1976), Tenney *et al.* (1992) proposed a cognitive framework in which

SA is the state of the 'perceptual cycle' at any given moment. They write *situation awareness can be thought of as the 'big picture' or context in which to interpret the flow of events. It allows the perceiver to attend to the right information at the right level of abstraction for the right task* (p. 3). The framework also includes anticipatory processes because, from Neisser's perceptual cycle theory, it is the anticipation of events that directs exploratory behaviour. To illustrate the importance of anticipation, Tenney *et al.* (1992) discuss the SA characteristics of an airline pilot whose job involves anticipating or 'thinking ahead of the aircraft' at all times. In routine situations anticipation is accomplished through the basic perceptual level; in non-routine and emergency situations, which require contingency planning and diagnosis, anticipation is achieved through a broader exploratory cycle based on knowledge of the world and consideration of possible outcomes. This notion of anticipation is central to the concept of SA and is discussed later (see 3.3).

Also focussing on perception, Finnie and Taylor's (1998) have proposed the 'Integrated Model of Perceived Awareness Control (IMPACT)' Model of SA. The model is a development of the 'Perceptual Control Theory (PCT)' of Powers (1973) and is an attempt to outline the cognitive processes involved in SA. The basis of this theory involves the concept that behaviour is associated with the control of perception. Furthermore, actions are only valuable if their outcome is favourably perceived in relation to the intended goals. Feedback is a fundamental requirement to goal directed behaviour, and this issue is a fundamental element of PCT. The PCT Model organises control under a hierarchical structure of multiple layers. Higher levels represent 'meta-goals' such as self-esteem and conscientiousness; lower level goals represent basic factors such as safety and survival. An action or behaviour is triggered in response to an error-correcting signal. The communicated signals aim is to change the environment so that the operator's perception meets this desired goal. Thus, PCT suggests that operator perception is controlled, not the behaviour.

Finnie and Taylor (1998) used this fundamental concept when developing the IMPACT Model of SA. According to this model SA, or the operator's perception of his or her own SA, is controlled by behaviour. Furthermore, the acquisition and maintenance of this SA is constructed from the behaviours involved in reducing the differences between the perceived level of SA, and the desired level of SA (see Figure 2).

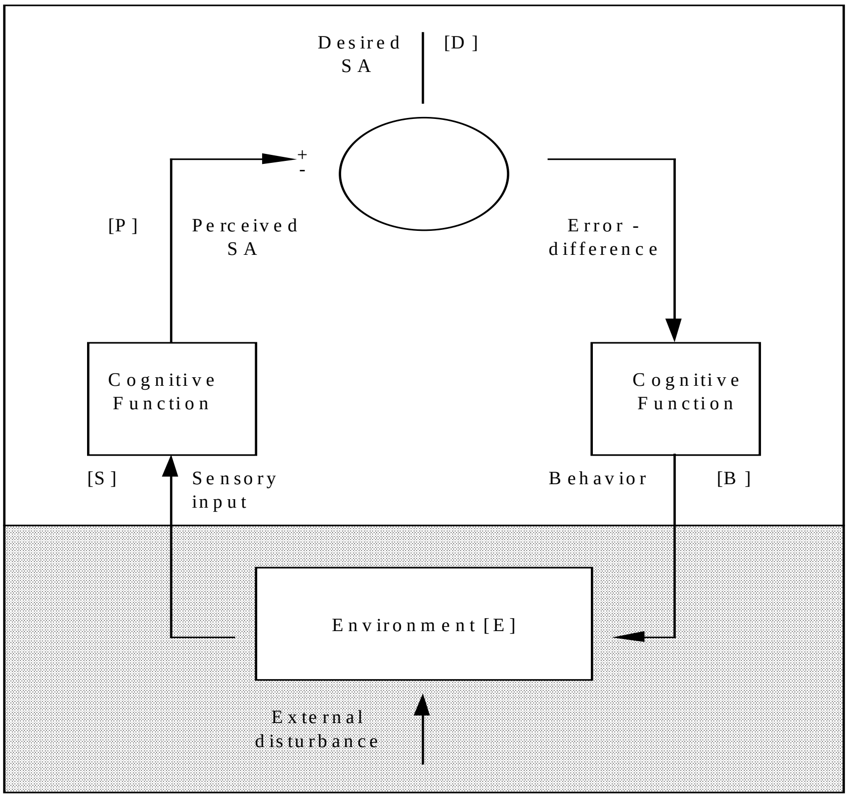

Figure 2: IMPACT Model of SA (Finnie & Taylor, 1998)

Page intentionally left blank

3. SITUATION AWARENESS AND ATM SYSTEMS

3.1 General Research

The first explicit consideration of controller's SA is to be found in the early nineties, in the wake of numerous studies looking at pilots' SA. However, if one considers SA as a form of mental representation, then the research origins can be traced back much further, for example to the studies of the controller's 'mental picture' carried out in the seventies. For the purposes of this report, the research can be grouped broadly into three categories:

1. Description of controller's SA and its main cognitive elements.

2. The relationship between controllers' errors and SA.

3. The study of the impact of automation on controller's SA. Within this third group, particular attention has been paid to the concept of free flight.

3.1.1 Cognitive elements of situation awareness

Garland *et al.* (1993) drew attention to the cognitive determinants of SA and emphasised the importance of working memory, cognitive skill acquisition, and automatic and controlled processing. The authors concluded with the statement: *Since the effects of advanced automation on ATC situational awareness are not yet known, it is imperative that direct manipulation workstation design be maintained to insure optimal situational awareness and performance* (p. 142).

Results of an experiment by Mogford (1994) with ATC trainees suggested that certain aircraft data is more critical than other data. He writes: *Although it might be expected that all aircraft information is critical for adequate air traffic controller SA, this experiment demonstrates that some elements (e.g. aircraft altitude and heading) may play a key role, whereas others (e.g. speed, position and identifier) may not be as important as expected... There may be three kinds of data in the ATC environment: a) that which must be remembered and updated; b) that which can be searched for when needed and forgotten; and c) that which can be ignored. Only the first type of data is retained in SA.* Interestingly, researchers developing measures of SA (e.g. SAGAT) have largely ignored this conclusion about the different levels of importance of SA elements.

More recently, researchers at the Technical University of Berlin (Niessen, Eyferth & Bierwagen, 1999) investigated the effects of controllers' experience upon their mental picture. They found that for the experienced controller, the picture is based on less, but more relevant, information compared to inexperienced controllers. At early stages of conflict detection the inexperienced controllers focus on every aircraft, whereas experienced controllers focus on the basis of particular features (i.e. a/c located near to each other, vertical movement or points in the space with higher probability of conflict). They classify the aircraft into two groups: those requiring further analysis and those that can be separated safely immediately.

As stated above, consideration of the controller's mental picture, which some researchers would argue is fundamentally the same concept as SA, has a long history (e.g. Bisseret, 1971; Whitfield & Jackson, 1982). Indeed, studies on the controllers' mental representation have been carried out since the late sixties, particularly in France. A synthesis can be found in Bisseret (1995). The main findings of these research studies, including related work by Ochanine (1969, 1981), can be summarised as follows:

- Controllers do not consider aircraft in isolation, but rather they consider aircraft in pairs of aircraft. The related data (SA elements) are thus relative data, e.g. 'this aircraft is at a higher level than this one' rather than 'individual levels'. The same point has been reiterated in more recent papers (Gronlund *et al.*, 1998).

- Controllers only consider the data needed to make decisions which can be thought of as a form of 'cognitive economy'. It was found that position and altitude are considered as two key data. It is only when information on position and altitude are not sufficient for conflict detection that the controllers look for other sources of information.

- Controllers operate in predictive mode. In a great majority of cases when the value of these two dimensions (position and altitude) are not reported correctly, it is because the controller was operating in a predictive mode, anticipating the situation (e.g. reporting position and altitude of aircraft two minutes ahead).

- Functional distortions have also been found in the representation of the airspace and map: these distortions are highly correlated with the frequency of traffic load on those elements.

For controllers 'having the picture' is the most important pre-requisite to carrying out their job, which is managing the safe, expeditious and orderly flow of their traffic. As remarked by Weston (1983), 'losing the picture' is the controller's nightmare.

3.1.2 Errors and situation awareness

The study of human error has for long been a subject of much research and, as mentioned in the Introduction, EUROCONTROL is very active in the area. The relationship between SA and error has been frequently studied. Several research studies have suggested that a controller's awareness of error development is related to the severity of operational errors. For example, Rodgers *et al.* (1995) studied a sample of 85 operational errors covering a three-year period. Characteristics of the operational errors were analysed with regard to the awareness or non-awareness of the controllers, and with regard to the complexity of the sectors and sector error rates. One of the main findings was that awareness of error development was significantly correlated with sector error rates; controllers involved in operational errors at sectors that have greater error occurrence tended to demonstrate less awareness of the developing error. The authors make the hypothesis that low awareness, presumably influenced by sector complexity, leads to higher operational error incidence. Similarly, Durso *et al.* (1997), in an analysis of data for a one-year period (1993), confirmed that controller awareness of error development resulted in significantly less severe errors.

The relationship between SA and error has also been investigated within the framework of Endsley's Model of SA (see 2.2 & 4.2), in particular related to the three levels of SA. According to Endsley's theory, 'level 1' SA errors are due to a failure to correctly perceive the situation. Specifically, the errors can be due to a lack of salience of a critical cue, a physical obstruction or a failure of the system to make the information available. They can also be caused by an over-abundance of information or lack of adequate strategy to direct the information sampling. Level 2 SA errors correspond to a failure to properly integrate or comprehend the meaning of perceived data in light of the operator’s goals. In a study by Jones and Endsley (1996), level 2 errors were attributed to an incomplete mental model, the use of an incorrect mental model and over-reliance on default value. Lastly, level 3 SA errors are due to over-projection of current trends, even in a situation that is clearly understood. However, as noted by Endsley (2000), although having a good level of SA does not a guarantee error-free performance, it is reasonable to suppose that having poor SA will increase the risk of errors occurring.

In a different vein, Endsley and Rodgers (1994) carried out a 'goal-directed' task analysis to identify SA requirements for en route controllers. The list is presented in a hierarchical format, listing first the controller’s main goal and associated sub-goals, and then each SA requirement for meeting these sub-goals. The detailed analysis does not, however, address the question of the information support, i.e. how a controller would actually get the required information. Neither are any weightings of importance attributed to the SA elements.

3.1.3 Automation and situation awareness

The effects of automation upon controller performance and particularly the controller's SA, is a topic of extensive research. Indeed, a conference dedicated to the topics of human performance, SA and automation was recently held[1]. The stated aim of this conference was to ‘focus on an integrated approach to system design that brings together a wide range of research on automation, information systems and situation awareness to focus on creating truly user-centred designs and advancing human/system performance’. However, it is beyond the scope of this report to mention more than a few relevant studies.

[1] See Web site: http://www.ie.msstate.edu/hpsaa/index.html.

Not surprisingly perhaps, Endsley and her colleagues have conducted a number of studies in this area (e.g. Endsley & Kiris, 1995a; Endsley & Kaber, 1999). Generally speaking, the results confirm that, even though full automation of task may be technically possible, it may not be desirable. Intermediate levels of automation may be preferable in order to keep human operators' SA at a higher level and allow him/her to perform critical functions. Similar recommendations have been made with regard to the introduction of free flight (Endsley, 1997; Endsley, Hansman & Farley, 1999).

3.2 Team Situation Awareness

As implied by the term, team situation awareness means the awareness held in common by team members. According to Prince and Salas (2000) *… at its simplest level, team SA is a construct that includes the individual SA of each team member and team processes* (p. 337). However, what is not so simple is determining if team SA refers to the total, combined SA of all the team members or only those elements of each individual's SA which are also shared or 'overlap' with other members' SA. As noted by Klimoski and Mohammed (1994), researchers are equivocal about what it is exactly that must be shared by team members. They write: *Most authors do not go much beyond describing (or ascribing) a collectivity of beliefs, shared understanding, or some similarity in the way information is processed. Conceptual issues regarding the content, form, and development of shared cognition remain largely unanswered* (p. 410).

Publications on team SA first appeared in the early nineties (e.g. Mosier & Chidester, 1991; Salas *et al.*, 1995) in the midst of much research into SA, team training (e.g. Swezey & Salas, 1992) and Crew Resource Management (CRM) (e.g. Wiener, Kanki & Helmreich, 1993). Indeed, most of the research into team SA has been within the context of CRM, or 'SA in crews' as Prince and Salas (*op. cit.*) refer to it. It focuses very much upon task performance and errors as a means of developing appropriate team behaviours and skills. This line of research remains very active (Prince & Salas, *op. cit.*; Salas, Muniz & Prince, 2000).

Other research has investigated team SA from the viewpoint of the team's mental models or, what has been called the 'shared' mental models. Cannon-Bowers and Salas (1990) proposed extending the concept of individual mental models to teams arguing that shared mental models should help to improve team performance because of the shared knowledge about the team and its objectives. Bolstad and Endsley (1999) have recently provided some experimental evidence for this proposal. However, as the definition of the shared mental model was that team members had been given information about their colleagues' tasks and an opportunity to discuss joint behaviour, the finding that this enhanced performance compared to the experimental condition of no sharing of information, is hardly surprising. Other researchers have looked at team SA as a form of 'distributed cognition' (Artman & Granlund, 1998; Artman & Garbis, 1998). Based on empirical data from two field studies in the domain of emergency control centres the research provides some interesting insights into the communication practices of teams.

In particular, the researchers conclude that understanding the interactions and consequences of the team's actions is more important than having an understanding (and mental model) of the system's dynamics. According to Artman and Granlund (*op. cit.*), *this does not change the need for adequate situation awareness but it might direct attention away from the situation of the system per se to the resources and practices needed to support coordination.*

Turning to the ATM domain it is evident that team SA in controllers has been somewhat sporadically investigated. The few studies that have been carried out have focussed upon the sharing of information between controller and pilot rather than between controllers (e.g. Javaux & Figarol, 1995; Farley *et al.*, 1998; Endsley, Hansman & Farley, 1999). This is perhaps not so surprising when one considers that the study of teamwork in ATC is relatively new. Although historically ATC has always been regarded as teamwork, it is also true that that controller functions are considered to be individual ones in terms of skills and accountability and most ATC training programmes are aimed at the individual controller (Ruitenberg, 1998). Indeed, the EUROCONTROL Team Resource Management (TRM) Programme, which is analogous to CRM in the aviation domain, treats SA solely in terms of the individual controller.

If the study of team SA, and especially team SA in ATC, has to date been incomplete, the study of the **measurement** of team SA is even more fragmentary. Kraiger and Wenzel (1997) proposed four measures of shared mental models based on their framework of 'antecedents, outcomes and components'. The proposed measures, processing information, organising information, shared attitudes and shared expectations, represent different aspects of the underlying cognitive components - the real measures are in fact well-known techniques such as questionnaires, protocol analysis, card-sorting and pairwise ratings. Based on an extensive review of available techniques, Langan-Fox, Code and Langfield-Smith (2000) also concluded that the four techniques that offer potential for the elicitation of a team mental model are cognitive interviewing, visual card sorting, causal mapping and pairwise ratings.

Regarding the observation of task performance, Langan-Fox *et al*. (*op. cit.*) do not recommend it. They write: *It is hard to envisage how it could be used to derive team mental models, as it would be difficult to determine shared understanding of a domain through observing behaviour. Observation of task performance is best suited to the examination of (individual) mental models in contexts for which user-system interaction is highly structured* (p. 252). However, a methodology for measuring team SA, which is heavily based on observation, has recently been developed by researchers at the US Naval Air Warfare Centre (Muniz *et al.*, 1998). The methodology is called 'Situational Awareness Linked Indicators Adapted to Novel Tasks (SALIANT)'. It is one of several measures discussed in more detail in 4.9).

3.3 The Controllers' View of Situation Awareness

Most of the literature on SA has been written by psychologists, human factors and other specialists who, although knowledgeable about ATC, are not air traffic controllers. Moreover, much of this research has been made in the context of aviation. It is essential therefore to get the controllers' point of view.

Fortunately, a number of operational (or ex-operational) controllers have published their views. Bert Ruitenberg, a controller assisting the International Civil Aviation Organization (ICAO) on the area of human factors, has proposed a descriptive model of situational awareness (Ruitenberg, 1997). According to Ruitenberg, many definitions of SA, for example that of Endsley, refer to 'elements of the situation' or 'elements of the environment' but fail to explain what exactly those elements are; to quote: 'Yet this is exactly where I feel her studies and those of several other scientists fall short: they fail to identify some of the elements that are crucial to us, Air Traffic Controllers'.

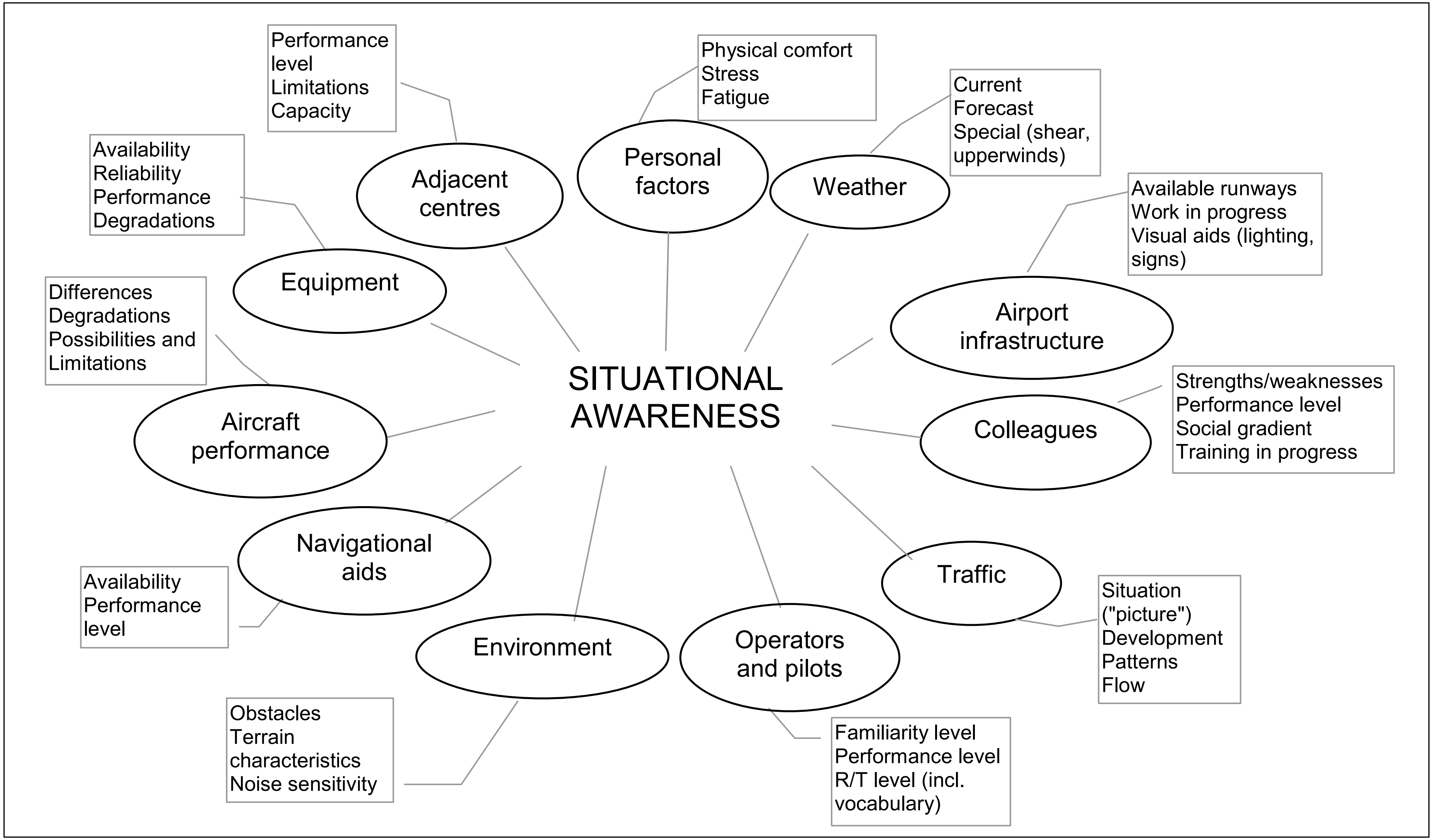

Figure 3: Elements of situational awareness (Ruitenberg, 1997)

Ruitenberg's Model (see Figure 3) is actually a description of all the elements that will or could influence the way a controller works at any particular moment, and in which subtle to large changes may occur at short notice. Ruitenberg emphasises the fact that *many of these changes can only be recognised after having gained considerable experience in ATC in general and at a specific location in particular*. It is interesting to note that 'traffic' is just one of the eleven elements identified. Thus, the controller's mental picture of the traffic is important, but is not the only constituent of his/her SA.

Controllers' views about SA have also been obtained directly from interviews with them. Jeannot (1999) carried out several structured interviews both at the EUROCONTROL Experimental Centre (EEC) and at the Athis-Mons Air Traffic Control Centre (ATCC). The observations and interviews at the Athis-Mons ATCC[2] were particularly interesting and fruitful for many reasons. First, there was the obvious advantage of observing real-life operations rather than simulations. Second, was the presence of trainees and On-the-Job Training Instructors (OJTIs). OJTIs are highly qualified controllers who are used to observe the behaviour of their trainees extremely carefully. Their objective is twofold: providing the best advice for the trainees to improve their performance but also, at the same time, making sure that the traffic is handled in a safe and expeditious way.

The interviews, made directly after the observations, were articulated around the following main points:

- What is situation awareness?
- What does it mean in their daily practice?
- Can you teach it?
- How do you know or realise that a trainee has, or has not, a good level of SA?
- What are the indicators that a trainee's SA is degrading?
- What are the strategies, if any, to regain SA?

The information gathered from the interviews provided a detailed description of controllers' views of SA. The key findings are summarised in Tables 1 & 2. In addition to the factors identified that affect SA, Jeannot (*op. cit.*) also notes the importance of the controller's experience on the job, his/her experience and knowledge of the sector, and his/her expectations from knowledge of previous, more or less similar situations.

[2] Nine controllers were interviewed: seven were from the Athis-Mons ATCC, two from Aix-en-Provence ATCC and six were also on-the-job trainers.

Table 1: Indicators of good and bad SA (Jeannot, 1999)

Indicators of good SA	Indicators of reduced/impaired SA
• Anticipating events • Being able to predict next a/c call • Managing resources (technical system , internal, team …) • Managing time • Feeling of being in control; able to implement elegant solutions • Taking the right decision at the best moment, managing traffic in a safe and expeditious way • Detecting mismatches	• Increase in delay between pilot calls and controller answer • Inconsistency in communication with pilots, colleagues, adjacent sectors or centre • Sudden and unexpected variation of workload • Confusion • Need to check same information several times

Table 2: Factors affecting SA and strategies to recover SA (Jeannot, 1999)

Factors leading to loss of SA	Strategies used to recover SA
• Time pressure • Focusing on non-pertinent, or less pertinent information • Focusing on a subset of relevant information but missing the evolution of other information • Becoming reactive rather than proactive • Reduction in room to manoeuvre, • Increased occurrence of non-safe situations • Noise/distraction other people • Mental and/or physical fatigue • Volume of traffic; unexpected and sudden variation of traffic load • Number of phone calls • Lack of strip information at the right time • Lack of adequate feedback • Too much happening and having to process too much information • Traffic building up • Unusual or unexpected events (e.g. aircraft calling in too early)	• Increase in delay between pilot calls and controller answer • Check consistency between strips and radar • Force themselves to speak slowly and precisely, return to strict phraseology • Request help: ask for sector splitting, if possible, to decrease load • Analyse closely all the strip • Force themselves not to spend too much time on a single problem • Change principle of strip classification, re- organise strips according to new criteria (e.g. by entry or exit point) • Physically manipulate the strips (i.e. re-positioning the strips aids concentration) • Prioritise work and “forget” less important tasks • Disregard the strips of ‘hello-goodbye’ a/c (in the case of too many strips) • Always prioritise new strips

Other controllers' views about SA were also gathered as part of the validation activity for the SHAPE SA Measure (see 5.4). The first part of the interview conducted with instructors at the EUROCONTROL Institute of Air Navigation Services (IANS) was on their perception of what exactly SA in the context of ATC. As well as echoing the controllers' views gathered previously, one instructor gave an interesting definition of SA as, 'what you need to know not to be surprised'.

Instructors emphasised the importance of the flight strip as an external support of SA. Time was spontaneously proposed as an indicator of good or bad SA: time to answer an aircraft call as well as time to find an information. The following indicators of poor or degrading SA were specifically mentioned:

– focusing on a problem inside the sector,
– forgetting to transfer aircraft,
– forgetting to give Exit or Requested Level,
– allowing aircraft to deviate from their route,
– beginning to be surprised.

It was underlined that recovering from a degradation or loss of SA was extremely difficult and could be done only by experienced controllers with an excellent knowledge of the sector.

The information gathered from these various interviews provides not only a good understanding of how controllers perceive SA, but also valuable insights about how to measure it as described later in the report (see Section 5).

Situation awareness (SA) from the controller's point of view is best summed up by the following quotation: "SA is what you need to know not to be surprised".

4. SITUATION AWARENESS MEASURES

4.1 Introduction

A number of measures of SA have been developed over the years. Each has its advantages and disadvantages, and some their well-publicised protagonists. The measures can be grouped in three categories:

1. **Query techniques**, in which the subjects are asked (‘queried’) directly about their perception of certain aspects of the situation.

2. **Rating techniques**, in which either the subjects themselves, or observers of the subjects, are asked to rate SA along a number of dimensions, typically presented in a series of scales.

3. **Performance-based techniques**, in which the level of SA is inferred from the level of performance. The rationale underlying this technique is that good SA is needed to achieve a good performance.

The goal of this chapter is to review the main existing measures of SA that have either been used in the context of ATC simulations, or appear to be most relevant to them. nine measures of SA have been identified as shown in Table 3. Five of the measures are query techniques (i.e. SAGAT, SPAM, SAVANT, SALSA and SAPS) and four are rating techniques (SART, C-SAS, SALIANT and SA/BARS). No performance-based techniques have been included.

Table 3: Existing situation awareness measurement techniques

	Situation Awareness Measures	
1	SAGAT	Situation Awareness Global Assessment Technique
2	SART	Situational Awareness Rating Technique
3	SPAM	Situation Present Assessment Method
4	SAVANT	Situation Awareness Verification Analysis Tool
5	SALSA	Situation Awareness bei Lotsen der Streckenflugkontrolle im kontext von Automatisierung.[3]
6	SAPS	Situation Awareness ProbeS
7	C-SAS	Cranfield Situation Awareness Scale
8	SALIANT	Situation Awareness Linked Indicators Adapted to Novel Tasks
9	SA/BARS	Situation Awareness Behaviourally Anchored Rating Scales

[3] Translated as ‘situation awareness of en-route air traffic controllers in the context of automation’.

For each of the nine measures, the underlying theory is presented, and the procedure to apply it is briefly described. The overall advantages and disadvantages are also summarised.

Finally, in 4.10 the implications for SHAPE of the review of the measures are analysed, including the suitability of applying each measure within the context of real-time simulations at the EEC. Possible adaptations and/or enhancements to the measures are also considered.

4.2 SAGAT

4.2.1 Theory

The Situation Awareness Global Assessment Technique (SAGAT) is a query technique developed by Endsley (1988, 1995b, 2000) and Endsley and Kiris (1995b). As described in 2.2, SAGAT is based on information-processing theory. SA is an internal model derived from the environment that is separate from, and precedes decision-making and performance.

SAGAT is the most widely known measure of SA and certainly the best publicised[4]. However, despite the numerous studies that have been conducted over a wide range of application areas, the validity of the technique remains debatable. For example, in a comparative simulation study (Endsley *et al.*, 2000), SAGAT was found to be no better than SART or a SPAM-like probe technique.

4.2.2 Method

The simulation is frozen at randomly selected times and subjects are queried as to their perception of the situation at that instant. SAGAT queries are on specific data or data criteria corresponding to the three levels of SA. The screen and all information sources are blanked/hidden. Computerised versions of SAGAT exist, but paper versions are probably more easy to use (and modify).

4.2.3 Data collected and scoring

The data that is collected corresponds to the three levels of SA depicted in Endsley's Model. The majority of the data collected corresponds to values of Level 1 data. For ATCOs the first query will always (i.e. at every simulation halt) address the position of the aircraft in the controlled sector. An appropriate sector map is presented and the subject is required to enter the location of all aircraft within the sector and the immediate area surrounding it.

Responses are scored, as correct or incorrect. Questions asked, but not answered, are considered incorrect. All aircraft are considered equivalent. For some data an 'acceptable tolerance band around the actual value' exists.

High SA in SAGAT means that the situation and the subjective image of the situation match to a high degree.

[4] See SA Technologies Web site: http://www.satechnologies.com.

4.2.4 Verdict

Advantages	Disadvantages
• Recognises the need for ‘a comprehensive assessment of operator SA requirements’ for the design of the queries • Quantitative results • Best known (and widely used) SA evaluation method = possibility to compare with similar data in similar context	• Interruption of the natural flow of the task • All aircraft are considered equal • Does not take into account the principle of operational distortion: controllers change (unconsciously) some aspect of reality to make it easier to work

4.3 SART

4.3.1 Theory

The Situational Awareness Rating Technique (SART) is a multi-dimensional rating technique developed by QinetiQ[5], formerly known as the UK Defence Evaluation and Research Agency (DERA) (Taylor, 1990, 1995a, 1995b; Taylor *et al.*, 1995). There are three primary SART rating dimensions, corresponding to the three clusters of the original constructs elicited from military aircrew:

- Demand on attentional resources (D),
- Supply of attentional resources (S),
- Understanding (U).

This simplified form of the technique is referred to as the 3-D SART. The original constructs also provide ten secondary rating dimensions nested within the three primary ones. This is referred to as the 10-D SART.

According to Jones (2000), numerous studies have been performed to examine the validity of SART. Strong claims are made for the validity and sensitivity of the scale constructs, and the diagnostic capability of SART, but the evidence remains weak at best. In a recent empirical ATC simulation study comparing three SA measures, SART was found **not** to be sensitive to the display manipulation (Endsley *et al.*, 2000).

4.3.2 Method

At the end of each exercise (or series of exercises) the subject is asked to rate their knowledge or understanding of the situation associated with the task performed. For each of the three or ten dimensions the subject provides a score using a seven-point Likert scale labelled Low (score 1) to High (score 7).

4.3.3 Data collected and scoring

SART was not developed to provide a unitary measure of SA, but an overall SA score can be obtained by using the following simple algorithm:

SA (calculated) = Understanding - (Demand -Supply)

[5] See Web site http://www.qinetiq.com.

4.3.4 Verdict

Advantages	Disadvantages
• Non intrusive, i.e. no interruption of the task flow • Ease of use • Can be used for the evaluation of real (i.e. already in operation) systems or situations (therefore, not only in the framework of a simulation)	• Highly subjective (self-rating): subject perception of his/her SA; can be influenced by perceived performance • Scales, dimensions are too generic and too far from the concrete activity, therefore not really informative for system design • Not adapted to specifics of particular tasks/jobs

4.4 SPAM

4.4.1 Theory

Researchers at the University of Oklahoma have developed the Situation Present Assessment Method (SPAM)[6]. It is based on the premise that SA involves simply knowing where in the environment to find a particular piece of information, as opposed to remembering what that piece of information is (Durso *et al.* 1995; 1998). For example, a controller need not store in memory the call sign of an aircraft, but good SA may require that he/she know where to find the call sign should communication with the aircraft be required.

An evaluation of the SPAM Technique was carried out at the EEC during 2000 as part of the EATCHIP Phase IIIa experiments. Preliminary results indicated that the SPAM Technique was promising, but not without a number of problems (Jeannot, 2000).

4.4.2 Method

The controller is asked questions via his/her landline. The simulation is not stopped and all the information remains available to the controller. When the controller answers the telephone call, the researcher reads the question and starts a timer.

4.4.3 Data collected and scoring

Response time for correct responses rather than proportion of correct responses. Actually, two response times are taken into account: the time to answer the landline call as an indicator of workload, and the time to answer the actual query as the indicator of SA. The queries ask for 'gist type' information (e.g. which of the two aircraft has the lower altitude?) rather than for unique data on a single aircraft. The data analysed is the time between the query and the correct answer.

[6] See Web site: http://www.ou.edu/HTIC.

4.4.4 Verdict

Advantages	Disadvantages
• Use of landline can be fitted into the controller's existing activity schema • Less intrusive than other query techniques • Uses "gist type" questions (some research evidence that ATCOs process information in this manner) • No freeze, no stop of the simulation • ATCO support (i.e. length of time taken to respond considered by ATCOs to be a good indicator of SA)	• Scripted questions mean it is difficult to use in multi-sector simulation • Importance of queries not rated • Questions and answers are given verbally (Spatial representation of the traffic is more difficult than when aided with a map as with SAGAT) • Speed of the response is dependent on workload and spare capacity

4.5 SAVANT

4.5.1 Theory

SAVANT is a technique developed by researchers at the FAA Technical Centre (Willems, 2000). SAVANT is not based on a new theory or model of SA, but rather is an amalgamation of two other query techniques, namely SAGAT and SPAM. The SAVANT Measure is an attempt to retain and combine the advantages of both the latter techniques. According to Willems (*op. cit.*), the specific advantages to be retained from SAGAT are:

- queries are 'anchored' in the airspace (i.e. the location of aircraft on the sector map);
- the controller enters his/her responses directly into the computer system.

From SPAM the specific advantages to be retained are:

- no interruption of the simulation,
- no extensive use of memory,
- queries of relational information instead of verbatim information.

The use of SAVANT was evaluated in an extensive simulation (Willems, Heiney & Endsley, 2001, in press). In another simulation study investigating team configurations, the results indicated that SAVANT was less useful than expected (Willems, 2001).

4.5.2 Method

Queries on aircraft pairs 'à la SPAM' and more global sector-based queries 'à la SAGAT' are asked during the simulation without it being halted. Concerned aircraft are highlighted on the screen to limit the need for visual search, only the data directly concerned by the queries is blanked.

4.5.3 Data collected and scoring

Responses to queries, correctness and response time.

For the analysis aircraft pairs queries 'à la SPAM' and sector-based queries 'à la SAGAT' are separated. For the aircraft pairs queries the response times are analysed for correct answers, and also for all answers. Another comparison is made between queries on the present situation and queries on the future situation.

For the sector-based queries the percentage of answered queries and the percentage of correctly answered queries is calculated. A Subject Matter Expert (SME) evaluates the correctness of the answers given.

4.5.4 Verdict

Advantages	Disadvantages
• No simulation halts • Queries on aircraft pairs • Response time recognised by ATCO themselves as a good SA indicator	• Response mode (need to grasp the mouse and move it into position to answer) • Preparation of the queries is more sensitive than for SPAM (affects interface) • Current setting, with answering display on a different screen has been judged as disruptive by ATCOs

4.6 SALSA

4.6.1 Theory

Researchers at the Technical University of Berlin[7] have developed the SALSA Measure of SA (Hauß, Gauss & Eyferth, 2000, 2001). The measure is basically a variant of SAGAT (see 4.2). In addition to the rate of correct reproduced elements of the task environment, SALSA pays special attention to the operational relevance of these elements.

4.6.2 Method

Simulation is interrupted (e.g. thirteen times in 45 minutes), the radar screen is frozen and a single aircraft is highlighted. During each interruption the complete set of thirteen questions is asked for one aircraft. SALSA includes an expert rating of the replay of the simulation to determine the relevance of each item that is asked in the reproduction test.

4.6.3 Data collected and scoring

Only correct reproduced items that were judged as relevant in the replay are considered.

4.6.4 Verdict

Advantages	Disadvantages
• Recognises the need for 'a comprehensive assessment of operator SA requirements' for the design of the queries • Recognise that relevance of the elements of the task environment changes with times and circumstances • Quantitative results	• Interruption of the natural flow of the task • The weighting procedure occurs after the simulation: only the relevance of answers is taken into account, not the relevance of questions

[7] See Web site: http://www.tu-berlin.de/eng/index.html.

4.7 SAPS

4.7.1 Theory

The SAPS Technique was originally developed to test objectively the SA of helicopter pilots (Deighton, 1997; Jensen, 1999). The overall purpose being to 'assessing the actual SA achievable in simulation and flight trials assessments of new systems' (Jensen, 1999). This technique involves questioning the pilot, using a series of single word, pre-prepared questions ('probes') that are rehearsed with the pilots prior to the test.

4.7.2 Method

First version: 'Retrospective' = queries and confidence ratings at the end of the mission,

Second version: 'Concurrent' = queries and confidence ratings at pre-determined points of the mission. In this version pilots were briefed and trained to expect the probe questions at the specified points along the mission. The mission is divided in sub-elements.

The probe questions are shortened to one or two words, so as to be as non-intrusive as possible. The pilot then responds with a single, simple reply.

4.7.3 Data collected and scoring

Data consists of replies to the coded probe questions, plus the pilot's reported confidence level in the answer given (i.e. confident, sure, not sure, guess).

4.7.4 Verdict

Advantages	Disadvantages
• Query on-line (no simulation halts) • Takes into account confidence levels • Low intrusiveness	• Scoring and interpretation of results • Memory may be subject to decay • Possible increase in workload • Pilot training • Start and end points of sub-elements difficult to determine and recognise • Risk of probes becoming primary task

4.8 C-SAS

4.8.1 Theory

The Cranfield Situation Awareness Scale (C-SAS) has been developed by Cranfield University[8]. The objective of the measure is to assess the SA of student pilots during their training. It is based on pilot actions and knowledge that the aviation community considers being important to maintaining SA (Dennehy, 1997).

An additional objective is actually to use this rating scale to promote changes in training by providing a satisfactory mean to monitor a pilot's progress in developing the skills necessary for SA.

4.8.2 Method

The C-SAS was designed for both in-flight/simulation and post-flight/simulation administration. The measure is employed either by the flight instructor (i.e. observer rating) or by the student pilot himself/herself (i.e. self-rating). Ratings may be obtained either during a pre-defined flight task, after task segments or following an entire task. The method entails providing numerical ratings for each of five sub-scales:

- pilot knowledge,
- understanding and anticipation of future events,
- management of stress, effort and commitment,
- capacity to perceive, attend, assimilate and assess information,
- overall situation awareness.

4.8.3 Data collected and scoring

There are two versions of the scale: a long and a short form. Ratings are collected for each of the five sub-scales. The overall SA score for each student is then calculated by simply adding all of the sub-scales scores together. A high score indicates a high level of SA, whilst a low score indicates a lower level of SA.

[8] See Web site: http://www.cranfield.ac.uk/coa.

4.8.4 Verdict

Advantages	Disadvantages
• Ease of use • Can be used for the evaluation of real (i.e. already in operation) systems or situations • Low intrusiveness especially if observer rating and/ or rating after the task • Sub-scales adapted to the specificity and characteristic of the task	• Subjective especially if used as self rating • Focused on subject evaluation therefore not really informative for system design • Level of intrusiveness raises in the case of self rating during the task • Would be more difficult to use in ATC, less observable behaviours

4.9 SALIANT

4.9.1 Theory

The Situational Awareness Linked Indicators Adapted to Novel tasks (SALIANT) Measure has been developed by the US Naval Air Warfare Centre[9] so as to provide a methodology for assessing team SA. As described by Muniz *et al.* (1998) the development of SALIANT involved five steps:

1. Delineation of behaviours theoretically linked to team SA (24 behaviours were identified, categorised into five clusters).
2. Development of scenario events.
3. Identification of specific, observable responses.
4. Development of a script.
5. Development of an observation form (re: step-three responses).

A preliminary study with aircrews to provide empirical validation of the measure has been reported by Bowers *et al.* (1998). The results, it is claimed, *indicated some success* (p. 12-5). Also, in a study by Fink and Major (2000) the SALIANT Measure, in comparison to SART and a SA probe technique, was judged to have the best psychometric properties and to be the 'most promising'.

4.9.2 Method

The method entails essentially repeating the five-step methodology for the particular domain and system of interest. In particular, the 'specific, observable responses' must be identified and an observation form created.

4.9.3 Data collected and scoring

The data collected is the presence or absence of the identified behavioural responses (labelled 'acceptable response') to a pre-determined event in a segment of the scenario.

[9] See Web site at: http://www.ntsc.navy.mil.

4.9.4 Verdict

Advantages	Disadvantages
• Finely adapted to the specific tasks • Observation form is easy to use • Good inter-rater reliability • One of the few methods looking at Team SA	• Process is time-consuming and labour intensive • Limited to highly observable team behaviours. • Need of an existing theoretical data base of behaviours linked to team SA

4.10 SA/BARS

4.10.1 Theory

Researchers at the University of Queensland developed the Situation Awareness Behaviourally Anchored Rating Scales (SA/BARS) Technique for Air Services Australia (Neal *et al.*, 1998). SA/BARS consists of seven questions, each of which is answered on a seven-point rating scale. Points 1, 4 and 7 of the scale are 'anchored' with a description of the expected controller behaviour (or lack of it). According to Boag, Neale and Neal (1999), these scales are used as part of the regular performance assessment protocol for Australian ATCOs.

Interestingly, according to Boag, Neale and Neal (*op. cit.*), the SA/BARS questions are divided into three groups corresponding to the three levels of SA as in SAGAT (see 4.2). That is, question 1 assesses Level 1 SA; questions 2 to five assess Level 2 SA; and questions 6 and 7 assess Level 3 SA. However, it is not clear if this grouping is fixed or variable.

4.10.2 Method

SA/BARS is designed to be used primarily by a controller rating the performance of another controller, for example after observing a training session. The observing controller is required to simply rate the subject controller's performance in answer to each of the SA/BARS questions.

Boag, Neale and Neal (*op. cit.*) report an experimental study in which controllers also used SA/BARS to rate their own performance. The statistical data comparing SA/BARS observed ratings to the self-ratings shows that the two versions are highly correlated.

4.10.3 Data collected and scoring

The average score for the seven SA/BARS questions is calculated. The higher the score (from 1 to 7), the better the controller's SA.

(In the Boag, Neale and Neal (*op. cit.*) study the SA scores were calculated for each of the three SA levels. However, it is not clear whether or not this is the normal, preferred scoring method or calculated especially for the purposes of the experiment.)

4.10.4 Verdict

Advantages	Disadvantages
• Explicit SA measure focused on ATC tasks • Rating scale uses 'behavioural anchors' • Relatively simple to use • Low intrusion • Can be used either by an observer or for self-rating	• Behavioural anchors might be difficult to discriminate between • Anchors for three points only might bias scores away from intermediate points • Questions too broad • Subjective (as with any questionnaire)

4.11 Implications for SHAPE

4.11.1 Introduction

Nine measures of SA have been reviewed in the above sections of the report. The next step is to determine which of these techniques, if any, could be used for SHAPE, either directly in their current form or possibly adapted in some manner. As has been shown, all of the measures have various advantages and disadvantages. However, it is not immediately obvious which factors are the more important than others and how one should choose between the techniques. The purpose of this section is to provide such an analysis.

4.11.2 SAGAT

First, considering SAGAT, which is the most well-known SA measure, it is the view of the authors of this report that it is *not* suitable for the purposes of SHAPE. To some readers this is perhaps a surprising statement to make. However, the disruption caused by halting the simulation (i.e. as required by SAGAT) is considered too great. In addition and possibly more serious is the fact that SAGAT is not designed for use in multi-sector real-time simulations (i.e. it relies on carefully scripted questions for a single sector).

Another problem with SAGAT is that all of the questions are all treated as being of equal importance which, in the view of authors, ignores an important aspect of SA. Finally, despite an abundance of studies employing SAGAT, the evidence that the technique actually provides useful data is very limited. Although a number of enhancements to SAGAT could be envisaged (and indeed the SALSA Technique is such an attempt) the conclusion remains it is not the right choice for SHAPE.

4.11.3 SALSA

For the same reasons advanced against SAGAT, the SALSA Technique is also not recommended. Although SALSA does address the problem of linking the SA questions to their operational importance, the technique suffers from too many practical problems (e.g. intrusive, not very easy to use). Moreover, concerning both SAGAT and SALSA, it should not be overlooked that the measures would probably interfere with other human factors measurements (e.g. workload) been taken.

4.11.4 SAVANT

Concerning SAVANT, the fact that it does not require halting of the simulation is clearly an improvement over SAGAT. However, the results from a recent simulation that employed SAVANT indicate that the technique proved less useful than expected (Willems, 2001). In particular, controllers found the mode of entering their responses (i.e. using a separate computer screen and input device) quite disruptive. It is also evident that the preparation time required for scripting questions is significant and from a purely practical point of view, not feasible for most 'ordinary' simulations conducted at the EEC. Moreover, even if resources for such preparation were readily available SAVANT could not be employed in multiple-sector simulations because scripted or planned-for events might not happen (i.e. because controller's actions in a preceding sector have changed the scenario). For these reasons, SAVANT, at least in its current form, is not recommended.

4.11.5 SPAM

As discussed earlier (see 4.5) the SAVANT Technique is also an attempt to utilise the best features of another SA measure called SPAM. Considering SPAM in its own right, it is the view of the authors of this report that it offers perhaps the most promising approach to measuring SA. Indeed, we are strongly in agreement with the developers of SPAM that, from a theoretical point of view, SA is not just about remembering data *per se*, but rather knowing where to find relevant information when needed. This also accords with the controllers' view of SA that it really about ‘what you need to know so as not to be surprised’(see 3.3). We also agree with Durso *et al.* (1995) who remarked *... if a piece of information was immediately available in the environment, it might be a poor idea to use limited resources to remember it.*

From a practical point of view, the fact that SPAM does not require the simulation to be halted (as SAGAT does) is a major advantage. In addition, asking the SA questions to the controller via his/her telephone landline minimises the level of disruption. As stated by Durso *et al.* (1995, 1998) responding in this manner fits well into the controllers' normal job. Regarding preparation time, which is a possible disadvantage of SPAM, one solution might be that questions are asked on-line by a Subject Matter Expert (SME), i.e. a controller. In fact, the involvement of SMEs in devising the SA questions, both their type and timing, seems a prerequisite for a true measure of SA. The implication of this proposal is that some facilities (hardware and software) would be required for the SME observers. Understandably, some projects might not have any interest in supporting such a sophisticated measure, at least not on a routine basis. This would seem to suggest that an alternative SA measure, such as a questionnaire, is necessary. Clearly, a questionnaire administered after an exercise is low-cost and non-intrusive.

4.11.6 SAPS

Regarding SAPS, although the principle of asking questions of operational relevance is good, the requirement to rehearse a limited set of questions limits the extent to which the technique could be applied to complex, dynamically changing environment such as ATC. Similarly, SAPS is also dependent on the division of the operator’s (pilot or controller) activities into clearly distinct phases that would be extremely difficult to identify for ATC. For these reasons SAPS is not recommended.

4.11.7 SART

Three questionnaire/rating approaches to measuring SA have been reviewed. Regarding SART, which like SAGAT is frequently cited in the literature, it is the view of the authors of this report that it also is **not** suitable for the purposes of SHAPE. The fundamental problem with the measure is that the questions that it uses are neither specific to task performance nor related to the task domain. In fact, in our view, the technique would not provide any meaningful insight into controllers' SA. This measure is therefore not recommended.

4.11.8 C-SAS

In contrast, the Cranfield Situation Awareness Scale (C-SAS) Measure suggests a relatively straightforward way to develop an ATC-related questionnaire. Although the C-SAS Measure is aimed at pilots rating other student pilots, the principle of asking a focussed, task-related set of questions could be used for SHAPE. That is, the questionnaire could be one that is completed by the controller himself/herself. The success of such an approach will clearly depend on devising a suitable set of questions.

4.11.9 SALIANT

The SALIANT Measure is particularly interesting because it focuses upon team SA. However, its dependence on rather elaborate scripting of scenarios and observable behaviours rules its out for SHAPE. That being said, the methodology is of relevance to the development of a measure of teamwork.

4.11.10 SA/BARS

The SA/BARS Measure is probably the best questionnaire-based SA measure currently available and the general approach is recommended. However, for the purposes of SHAPE several aspects would need to be changed. The behavioural anchors appear to be unnecessarily complex (especially with each having two behaviours to observe or to choose between).

In addition, rather than constructing the questionnaire around ‘rigid’ SAGAT-like questions, the questionnaire for SHAPE should be based on what controllers themselves consider to be important for SA (see 3.3). It is interesting to note that in the experiment reported by Boag, Neale and Neal (*op. cit.*), there was no correlation between SA/BARS and SAGAT scores!

4.11.11 Summary

The above review and analysis of the various SA measures are summarised in Table 4. In addition to listing the main advantages and disadvantages, the nine SA measures are explicitly compared to each other using a set of criteria (applicability to ATM systems, applicability to multi-sector real-time simulations, explicitness, intrusiveness, ease of use, preparation and cost) If the criterion is met, it is marked 'Yes'; if not, it is marked 'No'. If the criterion is partially met, it is marked with an 'approximately' symbol ('≈').

To conclude, it is the view of the authors of this report that the best option for SHAPE is to develop two new SA measures - one query-based, the other subjective. As noted by Jones (2000), *the strength in utilizing subjective assessment metrics is that of a complement rather than a replacement to other forms of SA metrics* (p. 127). The two proposed measures are:

- an adapted version of SPAM, presented to controllers on-line,
- a questionnaire, presented to controllers at the end of each simulation run.

The development of these two measures is described in Section 5.

Table 4: Comparison of existing situation awareness techniques

SA Tech	Advantages	Disadvantages	Evaluation against criteria		Overall verdict
SAGAT	▪ Explicit measure ▪ Broad 'global' testing of SA ▪ Moderately easy to use	▪ Disruptive simulation halts ▪ Memory intensive ▪ Low sensitivity ▪ Randomised questioning	▪ Applicable to ATM systems ▪ Applicable to multi-sector R/T simulation ▪ Explicit SA measure ▪ Focussed on real-time dynamic SA ▪ Low intrusion ▪ Easy to use ▪ Minimal preparation ▪ Low cost	Yes No Yes No No ≈ ≈ ≈	Not recommended
SART	▪ Simple to use ▪ Some diagnostic capability ▪ Not intrusive	▪ Questions too general ▪ Confounded with workload ▪ Very subjective	▪ Applicable to ATM systems ▪ Applicable to multi-sector R/T simulation ▪ Explicit SA measure ▪ Focussed on real-time dynamic SA ▪ Low intrusion ▪ Easy to use ▪ Minimal preparation ▪ Low cost	No ≈ No No Yes Yes Yes Yes	Not recommended
SPAM	▪ Explicit measure ▪ Queries of relational information ▪ Response time indicator	▪ Question preparation ▪ Intrusion on primary task ▪ Workload effects	▪ Applicable to ATM systems ▪ Applicable to multi-sector R/T simulation ▪ Explicit SA measure ▪ Focussed on real-time dynamic SA ▪ Low intrusion ▪ Easy to use ▪ Minimal preparation ▪ Low cost	Yes No Yes ≈ ≈ ≈ ≈ ≈	Recommended with modifications

Table 4: Comparison of existing situation awareness techniques (cont'd)

SA Tech	Advantages	Disadvantages	Evaluation against criteria		Overall verdict
SAVANT	▪ Explicit measure ▪ Queries of relational information ▪ Response time indicator	▪ Perceived as Disruptive ▪ Question preparation ▪ Intrusion on primary task ▪ Workload effects	▪ Applicable to ATM systems ▪ Applicable to multi-sector R/T simulation ▪ Explicit SA measure ▪ Focussed on real-time dynamic SA ▪ Low intrusion ▪ Easy to use ▪ Minimal preparation ▪ Low cost	Yes No Yes ≈ ≈ No No No	Not recommended
SALSA	▪ Explicit measure ▪ Expert ratings of replays ▪ Weighted results	▪ V. disruptive simulation halts ▪ One a/c at a time ▪ Weighting process	▪ Applicable to ATM systems ▪ Applicable to multi-sector R/T simulation ▪ Explicit SA measure ▪ Focussed on real-time dynamic SA ▪ Low intrusion ▪ Easy to use ▪ Minimal preparation ▪ Low cost	Yes No ≈ No No No ≈ ≈	Not recommended
SAPs	▪ No simulation halts ▪ Uses confidence levels ▪ Low intrusiveness	▪ Preparation ▪ Scoring of results ▪ Risk of probes becoming primary task	▪ Applicable to ATM systems ▪ Applicable to multi-sector R/T simulation ▪ Explicit SA measure ▪ Focussed on real-time dynamic SA ▪ Low intrusion ▪ Easy to use ▪ Minimal preparation ▪ Low cost	≈ ≈ ≈ No Yes ≈ No Yes	Not recommended

Table 4: Comparison of existing situation awareness techniques (cont'd)

SA Tech	Advantages	Disadvantages	Evaluation against criteria		Overall verdict
CSAS	▪ Simple to use ▪ Diagnostic capability ▪ Sub-scales adapted to the characteristics of the task	▪ Subjective ('ignorance is bliss') ▪ More difficult to use in ATC as a peer rating technique	▪ Applicable to ATM systems ▪ Applicable to multi-sector R/T simulation ▪ Explicit SA measure ▪ Focussed on real-time dynamic SA ▪ Low intrusion ▪ Easy to use ▪ Minimal preparation ▪ Low cost	No ≈ ≈ No Yes Yes Yes Yes	General approach is recommended
SALIANT	▪ Finely adapted to the specificity of the task ▪ Addresses team SA	▪ Time-consuming and labour intensive preparations ▪ First simulation to identify observable responses	▪ Applicable to ATM systems ▪ Applicable to multi-sector R/T simulation ▪ Explicit SA measure ▪ Focussed on real-time dynamic SA ▪ Low intrusion ▪ Easy to use ▪ Minimal preparation ▪ Low cost	≈ ≈ ≈ No Yes No No No	Not recommended
SA/BARS	▪ Explicit measure ▪ Uses 'behavioural anchors' ▪ Relatively easy to use ▪ Can be used by an observer or for self-rating	▪ Behavioural anchors too complex ▪ Anchors biasing scores away from intermediate points ▪ Questions too broad ▪ Subjective	▪ Applicable to ATM systems ▪ Applicable to multi-sector R/T simulation ▪ Explicit SA measure ▪ Focussed on real-time dynamic SA ▪ Low intrusion ▪ Easy to use ▪ Minimal preparation ▪ Low cost	Yes ≈ Yes No Yes ≈ Yes Yes	Approach is strongly recommended

Page intentionally left blank

5. SITUATION AWARENESS MEASURES FOR SHAPE

5.1 Introduction

Based on the review and analysis of SA measures (see 4.10), it was concluded that two measures should be developed for the purposes of SHAPE. The first measure is an on-line measure based on SPAM. The second measure is a questionnaire.

The two measures, which are described in the following sections of the report, are:

- SASHA_L (SA for SHAPE on-line),
- SASHA_Q (SA for SHAPE questionnaire).

These two measures do not, however, address team SA. It is felt that research into team SA is still in its infancy. Consequently it is perhaps a little premature to expect a measure of team SA to be developed when 'traditional' SA measures of individual operators not yet fully proven. That being said, it should be possible to measure team SA within the context of a broader measure of team processes.

5.2 SASHA On-Line

5.2.1 Design

SASHA on-Line (SASHA_L) is based on the SPAM Technique (see 4.4). Five modifications to the technique are proposed so as to enhance it for application in multiple-sector simulations. The modifications are:

1. A Subject Matter Expert (SME) views the controller's screen, including any Decision Support Tools (DST) on a supplementary display terminal, separate from the main simulation room (e.g. co-located with pseudo-pilots).

2. The queries are formulated by the SME in real-time taking into account the real scenario as it unfolds. In other words the SME asks a question when he/she decides it is pertinent to do so.

   In the original version of SPAM queries are prepared in advance using a preview of the scenario. In the context of a multiple-sectors real simulation, this is not possible. A single change in one of the previous sectors can actually have a dramatic effect on the expected situation, e.g. a 'planned' conflict does not occur anymore because of a change of level of one of the aircraft.

3. The queries are rated by the SME as to their operational importance. This is not included in the SPAM Technique.

4. The queries can directly address the use of particular tools or other aspects of the automation support. This is in contrast to SPAM that only focuses on the traffic.

5. The time for the controller to respond to the query is rated as either 'OK', 'too long', or 'too short', the latter indicating perhaps a lucky guess.

In the original SPAM Technique the time to answer the query was used as an indicator of SA. However, because of concern expressed by controllers as to the validity of this indicator (see 5.4.1), this feature was modified.

As with the original SPAM Technique, questions should not be on single items of data but on relationships between to or more items, e.g. relations between two aircraft, relation between one aircraft and the sector or a specific area. Questions should concern the future situation as well as the current situation; the former should be more numerous. Exactly how far into the future should be determined by the position/role of the controller: Clearly, it will be shorter term for the Tactical Controller (TC) and longer term for the Planning Controller (PC) (even longer for the Multi-sector Planning (MSP)). The Subject Matter Expert (SME) asking the questions should also take into account the extrapolation capability of the tools provided (e.g. Look Ahead Display in PD3). Examples of possible queries that might be asked with SASHA are shown in Table 5.

Table 5: Illustrative example of SASHA on-Line queries

No.	Time	Example SASHA on-Line queries
1	15:01	Will US Air 1650 and Continental 707 be in conflict if no further action is taken?
2	15:06	Which sector, shown in the communication < *tool* > window has requested a change of FL at handover?
3	15:10	Are there any speed-conflicts on the J74 airway?
4	15:18	What is the time of the situation displayed in < *tool* > window?
5	15:23	Are you expecting any significant increase of workload within the next < *15* > minutes?
6	15:29	Which aircraft needs to be transferred next?
7	15:35	Which aircraft has the faster ground speed, US Air 992 or Air France 2249?
8	15:40	Which of the two conflicts shown in < *tool* > is more critical?
9	15:46	Which aircraft would benefit from a direct route: BA1814 or AF5210?
10	15:51	Which aircraft is going to reach its Requested Flight Level first: AA369 or US Air 551?
11	15:54	With which sector do you need to coordinate AF222 Exit Level?
12	15:59	Which of the two conflicts shown in < *tool* > is more critical?

It is to be noted that in order not to disrupt the controller too much, queries should not be asked too frequently, perhaps a maximum of one query every five minutes. This also leaves the SME with sufficient time to complete the rating form (see below) during the course of the simulation. Ideally the SME should be provided with the technical means to replay the simulation exercise if he/she wishes, but the form can also be filled in during the simulation run it.

5.2.2 Implementation

First and foremost the SME, or SMEs, must be familiar with the ATM system being simulated. They must of course know about the airspace, route structures, and traffic, but also be fully familiar with the Human-Machine Interface (HMI) and especially the tools and other automation support.

The SMEs need to be trained to apply the SASHA_L Measure itself: what types of questions to ask, when to ask them, how to rate them and how to score the answers given by the controller. It should be noted that this also includes customising questions about decision support tools and other forms of automation, to the specific ATM system being simulated. Detailed instructions for using SASHA_L are provided in Appendix B.

To facilitate the recording and interpretation of the results, a form has been designed to enable the SME observer to write down the query and score both it and the controller's answer. The proposed form for SASHA_L is shown in Table 6. The form is also included in Appendix B with the instructions.

Table 6: SASHA on-Line form for SME

SME Rating Form			
Query:……………………………………………………………………………………… ……………………………………………………………………………………………… ……………………………………………………………………………………………… ………………………………………………………………………………………………			
Query's operational importance	- |___|___|___|___|___|___|___|___| +		
Answer's operational accuracy	Incorrect ☐	OK ☐	Correct ☐
Time to answer	Too Short ☐	OK ☐	Too Long ☐

5.3 SASHA Questionnaire

5.3.1 Design

The design of the SASHA Questionnaire (SASHA_Q) is based on well-recognised human factors principles, e.g. open/closed questions, vocabulary and clarity, rating scales, and so on (see Charlton, 1996).

The questionnaire incorporates both questions of a generic nature (i.e. not dependent or not related to specific aspects of the simulation) and questions aimed at particular tools, HMI or other features of the automation. As currently designed, there are six generic questions, three specific 'tool' questions, and one question addressing SA globally. The proposed questionnaire is shown in Appendix A.

Regarding the wording of the questions, they are based on information obtained directly from controllers concerning what they considered to be the most important indicators of SA (see 3.3).

5.3.2 Implementation

SASHA_Q is a self-rating questionnaire to be completed by the controllers involved in the simulation (i.e. it is not designed for other observers). It is intended that SASHA_Q would be given to each controller at the end of each simulation run.

Before the questionnaire is printed and made ready to give to the controllers, the questions relating to tools must be tailored to the specific system being simulated. This will be carried out by a human factors specialist or other SME.

Controllers will need to be briefed before completing the question. One briefing at the start of a simulation trial should be sufficient.

5.4 Validation Process

5.4.1 SASHA on-Line

As stated earlier, SASHA_L is based on the existing SPAM Measure of SA. As also previously described (see 4.4.1) the development of SPAM has undergone a number of empirical validation studies. For example, in the Durso *et al.* (1998) experimental study, SPAM was compared to other measures of SA. The results indicated that SPAM was successful in predicting the SME's evaluation of controller's performance, and also in predicting 'remaining-actions count'. In addition, results showed that response time (the primary dependent variable) was a good index of expertise and, according to the original assumptions, of SA. Overall the experts responded faster than the intermediates, who responded faster than the novices did. These differences were even greater for questions about future events. Experimental studies such as these provide some initial validation evidence for SASHA on-Line.

Regarding the intrusiveness of SASHA_L, it is expected to be low as has been found for SPAM. For example, in the study conducted at the Human Factors Laboratory of the FAA Technical Center (Endsley *et al.*, 2000) controllers subjectively rated the SPAM Technique intrusiveness as low. To the extent that the latter are valid, then the validity of the SASHA_L Measure is also valid too.

In addition to the research literature supporting the use of a SPAM-like approach, the validity of SASHA_L was evaluated by obtaining the views of three ATC training instructors at the Institute of Air Navigation Services (IANS) of EUROCONTROL, Luxembourg. The general comments on SA and its indicators were in agreement with the comments gathered during the Athis-Mons visit (see 3.3). It is interesting to note that time to respond to an aircraft call was cited by both instructors as a good indicator of SA.

- The principle of asking auditory questions to the controller while the simulation is still normally running was deemed perfectly acceptable. Indeed, answering question from fellow controllers, or from the aircraft pilot, is already part of the ATC job. The controllers did, however, express some concern about the nature of the query to be asked. One instructor confirmed that according to the situation, the probability of conflict or the traffic load, the question would not necessarily have the same importance each time it was asked. The same question might have different implications for SA if the aircraft are involved in a problem or do not require intervention any longer.

- The expertise of the SME formulating and asking the queries is clearly a critical issue. Both instructors insisted on the fact that the SME has to be an experienced civil controller, familiar with the airspace simulated and fully trained with the HMI and Tools used.

- The idea of providing the SME with a short evaluation form allowing him/her to rate the level of operational importance of the query and the operational accuracy of the response has been very positively judged.

- Concern was expressed about using response time as a measure because the response depends crucially on the nature of the situation and does not necessary reflect good or bad SA. In other words, a long response time does not automatically indicate poor SA. In order to avoid this problem it was proposed that the SME should simply rate the response time as either ‘too long’, ‘OK’ or ‘too short’. This recommendation has been incorporated in SASHA_L as shown in Table 6.

5.4.2 SASHA_Q

The proposed questionnaire measure of SA is a well-recognised subjective measurement technique. It has been developed in accordance with good human factors practice (e.g. Charlton, 1996) and to that extent is a valid approach. The benefit of using of a subjective measure of SA is not in doubt. Indeed, as Durso and Gronlund (2000) remarked *… applied research must consider subjective judgements of SA. Regardless of how good an operator's SA is, if the operator does not recognize the high level of SA he or she enjoys, mistrust and concerns about the system or situation will be an important part of the operator's job satisfaction and performance* (p. 286).

Clearly, it is also important to check the validity of the questionnaire with representatives of the ATC community. The face validity of SASHA_Q has been examined during in-depth interviews of ATC instructors at the EUROCONTROL Institute of Air Navigation Services (IANS). In addition, views of controllers were obtained informally by email. The views obtained were in general very positive and can be summarised as follows:

- The first global impressions of SASHA_Q were positive, even if the principle of a post-exercise questionnaire was considered "less exciting" than the query on line technique.

- The wording of the questions was liked because it used everyday vocabulary and was thus straightforward and easy to understand. One instructor said it was a good thing to use the word "feeling" - the questions are more personal and at the same time less judgmental.

- Controllers said that 'being surprised' (see question 3 in Appendix A) was an important indicator of SA degradation. One of the instructors thought, nevertheless, that controllers might have a tendency to deny this fact and to answer positively whatever happened.

- The question about forgetting to transfer an aircraft (see question 5 in Appendix A) was considered a good question and a good indicator of SA (confirming earlier findings). It was thought that controllers would respond honestly. A controller proposed another rating scale representing numbers of aircraft not transferred on time.

- The term information needs to be clarified. If the controller is repeatedly checking information that is dynamic and changing, this is normal practice that does not tell us anything about his/her SA. If, on the other hand, he/she is checking several times the same static information then this can be an indicator of difficulty to maintain SA. It has also been noticed that the fact of checking several times an item of information can be related to a lack of trust in the tool. In this case you are then monitoring the correct functioning of the tool rather than the information.

- Questions 8 to 10 were judged to be valid, but the majority of controllers and instructors emphasised that the answers to these questions would be highly dependent on the nature of the automation and/or new HMI tested in each simulation.

- Finally, the global rating of SA (see question 10 in Appendix A) was thought to be a good way to finish the questionnaire. Interestingly, some controllers[10] found the question very similar to the very first question, i.e. 'Be ahead of the traffic' was for them equivalent to 'Having a good Situation Awareness'.

5.4.3 Comparison of SASHA with other SA measures

To conclude this section on the validation of SASHA it is interesting to rate the two SASHA Measures against the same criteria that were used earlier to compare the nine SA measures (see Table 4). The ratings of SASHA_L and SASHA_Q are shown in Table 7.

So as to illustrate further the claimed advantages of SASHA a simple numerical scoring scheme has been used to compare all of the SA measures: 'Yes' equals 1.0, '≈' equals 0.5, and 'No' equals 0.0. Thus, the higher then score the better the measure.

In addition, as the eight criteria are not all equally important, a weighting factor has also been used. As shown in Table 7, the most important criteria have been given a weighting of 'x3' and the least important a weighting of 'x1'.

On the basis of these criteria, in the specific context of SHAPE, it can be seen that the SASHA on-Line and SASHA Questionnaire Measures are both rated the best. It should be noted, that in a different context with different criteria, or differently weighted criteria, other SA measures might be more appropriate.

[10] Aix-en-Provence ATCC.

Table 7: Comparison of SASHA with other SA measures against specific (SHAPE) criteria

	Applicable to ATM systems	**Applicable to multi-sector real-time simulation**	**Explicit SA measure**	**Real-time dynamic SA**	**Low intrusion**	**Easy to use**	**Minimal preparation**	**Low cost**	**Overall score**
Weighting	*x3*	*x2*	*x2*	*x3*	*x2*	*x1*	*x1*	*x1*	
SALSA	Yes	No	≈	No	No	No	≈	≈	5.0
SALIANT	≈	≈	≈	No	Yes	No	No	No	5.5
SART	No	≈	No	No	Yes	Yes	Yes	Yes	6.0
SAGAT	Yes	No	Yes	No	No	≈	≈	≈	6.5
SAPs	≈	≈	≈	No	Yes	≈	No	Yes	7.0
CSAS	No	≈	≈	No	Yes	Yes	Yes	Yes	7.0
SAVANT	Yes	No	Yes	≈	≈	No	No	No	7.5
SPAM	Yes	No	Yes	≈	≈	≈	≈	≈	9.0
SA/BARS	Yes	≈	Yes	No	Yes	≈	Yes	Yes	10.5
SASHA Questionnaire	Yes	≈	Yes	No	Yes	Yes	Yes	Yes	11.0
SASHA on-Line	Yes	Yes	Yes	Yes	≈	≈	≈	≈	12.5

6. CONCLUSIONS

1. This report has described the development and (partial) validation of two measures of situation awareness (SA) for use during ATM real-time simulation. Both measures are informed by extensive research literature, which has been reviewed.

2. The first measure, called 'Situation Awareness for SHAPE on-Line (SASHA_L)' is based directly on an existing SA measure developed by researchers at the University of Oklahoma. For the purposes of SHAPE this latter technique was modified in a number of novel ways.

3. The second measure, called 'Situation Awareness for SHAPE Questionnaire (SASHA_Q)' is a post-exercise self-rating technique. It consists of ten questions that were especially designed by taking into account the views of controllers themselves about SA and its indicators.

4. Perhaps surprisingly, two well-known and well-documented measures of situation awareness - SAGAT and SART - have **not** been recommended for use by EUROCONTROL. Indeed, based on the review carried out, several other SA measures (excluding SASHA) are better suited to the purposes of the SHAPE Project.

5. During the course of the work reported, it has not proved possible to validate the two proposed measures in real-time simulations, for example at the EEC. However, a number of controllers, ATC instructors and other researchers were consulted. In particular, controllers from different countries (Sweden, UK, France) provided useful comments on the proposed measures. Both measures were favourably judged.

6. In term of guidance for usage, attention is drawn to the importance of involving Subject Matter Expert(s) [SME(s)] for implementing the on-line measure. The accuracy of the method is dependent on the quality of the queries by the SME to the controllers. It is therefore essential that the SME receive proper training, not only on the method itself, but also on each particular simulation setting, operational concept, HMI and new tools.

7. The SA measures proposed for SHAPE are aimed at the individual controller. Although research on team SA and its measurement is being carried out, the research is still in its infancy. That being said, it should be possible to measure team SA as part of a broader measure of teamwork and team processes.

Page intentionally left blank

7. RECOMMENDATIONS

1. The situation awareness (SA) questionnaire, SASHA_Q, was developed in consultation with a number of air traffic controllers and instructors. Whilst this consultation exercise gives a good degree of confidence in the validity of the questionnaire, further consultation is clearly desirable. It is recommended that the questionnaire should be distributed to a wider audience via EUROCONTROL.

2. The SA questionnaire should be tested at the EEC during real-time simulations.

3. The requirements for implementing the proposed on-line measure of SA, (SASHA_L) should be further investigated in consultation with EUROCONTROL. In particular, representatives of the Human Factors Laboratory at the EEC should be consulted.

4. The validity of the proposed on-line measure (SASHA_L) should be evaluated during real-time simulations conducted at the EEC or elsewhere.

5. Contact with other researchers working on situation awareness measures, should be maintained. In particular, the Ben Willems at the FAA Technical Centre, and Frank Durso at the University of Texas.

Page intentionally left blank

REFERENCES

Artman, H. & Garbis, C. (1998). Situation awareness as distributed cognition. Paper presented at the *Ninth European Conference on Cognitive Ergonomics (ECCE-9)*, Limerick, 24th-26th August. EACE.

Artman, H. & Granlund, R. (1998). Team situation awareness using graphical or textual databases in dynamic decision-making. Paper presented at the *Ninth European Conference on Cognitive Ergonomics (ECCE-9)*, Limerick, 24th-26th August. EACE.

Bisseret, A. (1971). Analysis of mental processes involved in Air Traffic Control. *Ergonomics*, 14, (5).

Bisseret, A. (1995). *Représentation et décision experte; psychologie cognitive de la décision chez les aiguilleurs du ciel*. Editions Octares.

Boag, C., Neale, M. & Neal, A. (1999). Measuring situation awareness: A comparison of alternative measurement techniques. In: *Proceedings of the 10th International Symposium on Aviation Psychology,* pp 1240-1246.

Bolstad, C.A. & Endsley, M.R. (1999). Shared mental models and shared displays: An empirical evaluation of team performance. In: *Proceedings of the 43rd meeting of the Human Factors & Ergonomics Society*.

Bowers, C., Weaver, J., Barnett, J. & Stout, R. (1998). Empirical validation of the SALIANT Methodology. Paper presented at the *RTO HFM Symposium on Collaborative Crew Performance in Complex Operational Systems*, Edinburgh, 20-22 April. RTO/NATO.

Cannon-Bowers, J.A. & Salas, E. (1990). *Cognitive psychology and team training; Shared mental models in complex systems*. Paper presented at the Fifth Annual Conference of the Society for Industrial Organizational Psychology, Miami, Florida.

Carroll, J.M. & Olson, J.R. (1987). *Mental models in human-computer interaction. Research issues about what the user of software needs to know*. Washington D.C.: National Academy Press.

Charlton, S.G. (1996). Questionnaire techniques for test and evaluation. In: T.G. O'Brien & S.G. Charlton (Eds), *Handbook of Human Factors Testing and Evaluation*, 81-99. Lawrence Erlbaum Associates

Deighton, C.D.B. (1997). Towards the development of an integrated human factors and engineering evaluation methodology for rotorcraft D/NAW Systems, *QinetiQ Report No. DERA/AS/FMC/CR97629/1.0*, Dec. Farnborough: QinetiQ Ltd.

Dennehy, K. (1997). *Cranfield - Situation Awareness Scale, User Manual*. Applied Psychology Unit, College of Aeronautics, Cranfield University, COA report N0 9702, Bedford, January.

Dominguez C., Vidulich, M., Vogel, E. & McMillan, G. (1994). Situation Awareness. papers and Annotated bibliography (U). Armstrong Laboratory, Human System Center, ref. AL/CF-TR-1994-0085.

Durso, F.T. & Gronlund, S.D. (2000). Situation awareness. In: F.T. Durso *et al.*, *Handbook of Applied Cognition*, 283-314. John Wiley & Sons.

Durso, F.T., Truitt, T.R., Hackworth, C.A., Crutchfield, J.M., Nikolic, D., Moertl, P.M., Ohrt, D. & Manning, C.A. (1995). Expertise and Chess: a Pilot Study Comparing Situation Awareness Methodologies. In: D.J. Garland & M. Endsley (Eds), *Experimental Analysis and Measurement of Situation Awareness.* Embry-Riddle Aeronautical University Press.

Durso, F.T., Truitt, T.R., Hackworth, C.A., Crutchfield, J.M. & Manning, C.A. (1997). En Route Operational Errors and Situation Awareness. *The International Journal of Aviation Psychology*, 8 (2), 177-194.

Durso, F.T., Hackworth, C.A., Truitt, T.R., Crutchfield, J.M., Nikolic, D. & Manning, C.A. (1998). Situation Awareness as a Predictor of Performance for En Route Air Traffic Controllers*. Air Traffic Control Quarterly*, Vol. 6 (1), 1-20.

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

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

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

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

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

Edwards, E. & Lees, F.P. (1974). *The Human Operator in Process Control*, London: Taylor & Francis.

Endsley, M.R. (1995a). Toward a theory of situation awareness in dynamic systems. *Human Factors*, Vol. 37/1, 32-64.

Endsley, M.R. (1995b). Measurement of situation awareness in dynamic systems. *Human Factors*, Vol. 37/1, 65-84.

Endsley, M.R. (1997). *Situation awareness, automation and free flight*. Presentation to the FAA/EUROCONTROL Air Traffic Management R&D Seminar, June 1997, Saclay, France. http://atm-seminar-97.eurocontrol.fr/endsley.htm.

Endsley, M.R. (1988). Design and evaluation for situation awareness enhancement. In: Proceedings of the Human Factors Society 32nd Annual Meeting (Vol. 1), 97-101. Santa Monica: Human Factors Society.

Endsley, M.R. (2000). Theoretical underpinnings of situation awareness: A critical review. In: M.R. Endsley & D.J. Garland (Eds), *Situation Awareness Analysis and Measurement.* Lawrence Erlbaum Associates.

Endsley, M.R. & Garland, D.J. (2000). *Situation Awareness Analysis and Measurement*. Mahwah NJ: Lawrence Erlbaum Associates.

Endsley, M.R., & Kaber, D.B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, 42(3), 462-492.

Endsley, M.R. & Kiris, E.O. (1995a). The Out-of-the-loop Performance Problem and Level of Control in Automation. *Human Factors*, (37) 2, 381-394.

Endsley, M.R. & Kiris, E.O. (1995b). Situation Awareness Global Assessment Technique (SAGAT). TRACON Air Traffic Control Version, user guide, TTU-IE-95-02, Texas Tech University.

Endsley, M.R. & Rodgers, M.D. (1994). *Situation awareness information requirements for en route air traffic control*. Report DOT/FAA/AM-94/27. Federal Aviation Administration.

Endsley, M.R., Hansman, R.J. & Farley, T. (1999). *Shared situation awareness in the flight deck-ATC system*. Paper presented at the 1999 Digital Aviation Systems Conference, Seattle, Washington.

Endsley, M.R., Sollenberger, R., Nakata, A. & Stein, E. (2000). Situation Awareness in Air Traffic Control: Enhanced Displays for Advanced Operations. Technical note DOT/FAA/CT-TN00/01, Federal Aviation Administration, W. Hughes Technical Center.

Farley, T., Hansman, J., Endsley, M.R., Amonlirdviman, K. & Vigeant-Langlois, L. (1998). The effect of shared situation information on pilot/controller situation awareness and re-route negotiation. Paper presented at the *2nd USA/Europe ATM R&D Seminar*, Orlando, Florida, 1st-4th December.

Fink, A. & Major, D. (2000). Measuring situation awareness: A comparison of three techniques. In: Proceedings *Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium*, Savannah, Georgia, October 15-19.

Finnie, S. & Taylor, R. (1998). The Cognitive Cockpit. *Flight Deck International*, UK & International Press.

Garland, D.J., Stein, E.S., Wise, J.A. & Blanchard, J.W. (1993). Situational awareness in air traffic control: A critical yet neglected phenomenon. In: D.J. Garland & J.A. Wise (Eds), *Human Factors and Advanced Aviation Technologies,* pp 123-145. Embry-Riddle Aeronautical University Press.

Gentner, D. & Stevens, A. (1983). *Mental Models*. LEA Associates, Inc.

Gronlund, S. D., Ohrt, D., Dougherty, M., Perry, J.L. & Manning C.A. (1998). Aircraft importance and its relevance to situation awareness. Report DOT/FAA/AM-98/16. Federal Aviation Administration.

Hauß, Y., Gauss, B. & Eyferth, K. (2000). The Evaluation of a Future Air Traffic Management: Towards a new approach to measure Situation Awareness in Air Traffic Control. In: L.M. Camarinha-Matos, H. Afsarmanesh & H. Erbe (Eds), *Advances in Networked Enterprises: Virtual Organizations, Balanced Automation and System Integration*, CD Suppl., Boston: Kluwer.

Hauß, Y., Gauss, B. & Eyferth, K. (2001). The influence of multi-sector-planning on the controllers' mental models. In: D. Harris (Ed), *Engineering Psychology and Cognitive Ergonomics*. Vol. 5, 203-209. Ashgate Publishing.

Javaux, D. & Figarol, S. (1995). Distributed situation awareness: A concept to cope with the challenge of tomorrow. In: R. Fuller, N. Johnston & N. McDonald (Eds), *Human Factors in Aviation Operations*, 293-298. Aldershot: Avebury Aviation.

Jeannot, E. (1999). Situation Awareness: Recommendations to Project Leaders, *EEC Draft Note.* Brétigny-sur-Orge, France: EUROCONTROL Experimental Centre.

Jeannot, E. (2000a). Situation Awareness: Synthesis of Literature Search. *EEC Note No. 16/00, Project ASA-Z-EC.* Brétigny-sur-Orge, France: EUROCONTROL Experimental Centre.

Jeannot, E. (2000b). SPAM evaluation. EEC Note. Brétigny-sur-Orge, France: EUROCONTROL Experimental Centre.

Jensen, S. (1999). Perceived versus real situation awareness: towards more objective assessment of SA. In: L. Straker & C. Pollack (Eds), *CD-ROM Proceedings of CybErg 1999: The Second International*

Cyberspace Conference on Ergonomics, 327-334. Perth, Australia: International Ergonomics Association Press.

Jones, D.G. (2000). Subjective measures of situation awareness. In: M.R. Endsley & D.J. Garland (Eds), *Situation Awareness Analysis and Measurement,* 113-128. Mahwah NJ: Lawrence Erlbaum Associates.

Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space and Environmental Medicine*, 67(6), 507-512.

Klimoski, R.J. & Mohammed, S. (1994). Team mental models: construct or metaphor? *Journal of Management*, 20, 2, 403-437.

Kraiger, K. & Wenzel, L.H. (1997). Conceptual Development and Empirical Evaluation of Measures of Shared Mental Models as Indicators of Team Effectiveness. In: M.T. Brannick, E. Salas & C. Prince (Eds), *Team performance assessment and measurement.* Mahwah NJ: Lawrence Erlbaum Associates.

Langan-Fox, J., Code, S. & Langfield-Smith, K. (2000). Team mental models: Techniques, methods, and analytic approaches. *Human Factors*, 42, 2, 242-271.

Mogford, R.H. (1994). Mental Models and Situation Awareness in Air Traffic Control. *The International Journal of Aviation Psychology*, 7 (4).

Mogford, R.H. (1997). Mental Models and Situation Awareness in Air Traffic Control. *The International Journal of Aviation Psychology*, 7, (4), 331-342.

Mosier, K.L. & Chidester, T.R. (1991). Situation assessment and situation awareness in a team setting. In: *Proceedings of the Special Session on Situational Awareness at the 11th Congress of the International Ergonomics Association*, Paris, France, 17th July 1991.

Muniz, E.J., Stout, R.J., Bowers, C.A. & Salas, E. (1998). A methodology for measuring team situational awareness: Situational Awareness Linked Indicators Adapted to Novel tasks (SALIANT). Paper presented at the *RTO HFM Symposium on Collaborative Crew Performance in Complex Operational Systems*, Edinburgh 20-22 April. RTO/NATO.

Neal, A., Griffin, M., Paterson, J. & Bordia, P. (1998). *Human factors issues: Performance management transition to a CNS/ATM environment. (Final Report: Air Services Australia).* Brisbane: University of Queensland.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology.* San Francisco: W.H. Freeman.

Niessen, C., Eyferth, K. & Bierwagen, T. (1999). Modelling cognitive Processes of Experienced Air Traffic Controller. *Ergonomics,* Vol. 42 (11), 1507-1520.

Ochanine, D. (1969). *Rôle de l'image opérative dans la saisie du contenu informationnel des signaux*. Questions de psychologie.

Ochanine, D. (1981). *L'image opérative*. Actes d'un séminaire et recueil d'articles, Centre d'Education Permanente, Département d'Ergonomie et d'Ecologie Humaine, Université Paris I.

Powers, W.T. (1973). *Behavior: The control of perception*. Chicago: Aldine.

Prince, C. & Salas, E. (2000). Team situation awareness, errors, and crew resource management: Research integration for training guidance. In: M.R. Endsley & D.J. Garland (Eds), *Situation Awareness Analysis and Measurement.* Mahwah NJ: Lawrence Erlbaum Associates.

Rodgers, M.D., Mogford, R.H. & Mogford, L.S. (1995). Air traffic controller awareness of operational error development. In: D.J. Garland & M.R. Endsley (Eds), *Experimental Analysis and Measurement of Situation Awareness,* pp. 171-176. Embry-Riddle Aeronautical University Press.

Ruitenberg, B. (1997). Situational awareness in ATC: a model. *The Controller*, Vol 36, No.1.

Ruitenberg, B. (1998). Teamwork for air traffic controllers. In: *Proceedings of the Second EUROCONTROL Human Factors Workshop - Teamwork in Air Traffic Services,* EATCHIP Human Resources Team, HUM.ET1.ST13.0000-REP-02, Ed. 1.0, Released Issue. Brussels: EUROCONTROL.

Salas, E., Muniz, E.J. & Prince, C. (2000). Situation awareness in teams. In: *International Encyclopedia of Ergonomics and Human Factors,* 555-557. Taylor & Francis.

Salas, E., Prince, C., Baker, D.P. & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human Factors*, 37, 1, 123-136.

Smith, K. & Hancock, P.A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors*, 37, 1, 137-148.

Swezey, R.W. & Salas, E. (1992). *Teams: their training and performance*. Norwood, N.J.: Ablex Publishing.

Taylor, R.M. (1990). Situational Awareness Rating Technique (SART): The development of a toll for aircrew systems design. In: *AGARD Conference Proceedings No 478*, *Situational Awareness in*

Aerospace Operations. Aerospace Medical Panel Symposium, Copenhagen, 2nd-6th October 1989.

Taylor, R.M. (1995a). CC-SART: The Development of an Experiential Measure of Cognitive Compatibility in System Design. *Report to TTCP UTP-7 Human Factors in Aircraft Environments, Annual Meeting,* DCIEM, Toronto, 12th-16th June 1995.

Taylor, R.M. (1995b). Experiential Measures: Performance-Based Self Ratings of Situational Awareness. *International Conference on Experimental Analysis and Measurement of Situation Awareness*, Daytona Beach, 1st-3rd November 1995.

Taylor; R.M., Shadrake, R.H. & Bunting, A. (1995). Situational awareness, trust and compatibility: Using cognitive mapping techniques to investigate the relationships between important cognitive system variables. In: *Proceedings of the 79th NATO AGARD Aerospace Medical Panel Symposium on Situational Awareness: Limitations and Enhancements in the Aviation Environment*, pp. 7-1 to 7-14. Etat-Major de la Force aérienne, Brussels, Belgium, 24-28 April.

Tenney, Y.J., Adams, M.J., Pew, R.W., Huggins, A. & Rogers, W.H. (1992). *A principled approach to the measurement of situation awareness in commercial aviation.* NASA Contractor Report 4451, July.

Weston, R. (1983). Human factors in air traffic control. *The International Journal of Aviation Safety,* 1, 94-104.

Whitfield, D. & Jackson, A. (1982). The Air Traffic Controller's picture as an example of a Mental Model. In: G. Johannsen & J.E. Rijnsdorp, *Analysis, Design and Evaluation of Man-Machine Systems, Proceedings of IFAC Conference*, Baden-Baden, Germany.

Wiener, E.L., Kanki, B.G. & Helmreich, R.L. (Eds) (1993). *Cockpit resource management*. Academic Press.

Willems, B. (2000). Development of the Situation Awareness Verification and Analysis Tool (SAVANT). FAA William J. Hughes Technical Center. *Unpublished*.

Willems, B. (2001). Study of an ATC Baseline for the Evaluation of Team-configurations (SABET). FAA William J. Hughes Technical Center. *Unpublished*.

Willems, B., Heiney, M. & Endsley, M.R. (2001, in press). *Decision Support Automation Research in the En Route Air Traffic Control Environment: Interim Report*. Draft version, FAA Technical Center, Human Factors Laboratory, Atlantic City.

Page intentionally left blank

FURTHER READING

Brickman, B., Hettinger, L., Stautberg, D., Haas, M., Vidulich, M. & Shaw, R. (1998). The Global Implicit Measurement of Situation Awareness: Implications for Design and Adaptive Interface Technologies. In: M. Scerbo & M. Mouloua (Eds), *Automation Technology and Human Performance: Current Research and Trends*, 160-170.

Durso, F.T, Crutchfield, J.M. & Balsakes, P.J. (2001, in press). *Cognition in a Dynamic Environment*, University of Oklahoma, Department of Psychology, Norman.

Pritchett A.R., Hansman R.J. & Johnson E.N. (1996). *Use of testable responses for performance-based measurement of situation awareness.* Conference on Experimental Analysis and Measurement of Situation Awareness, Daytona Beach.

Sarter, N. & Woods, D. (1991). Situation Awareness: A Critical but ill-defined phenomenon. *The International Journal of Aviation Psychology,* 1 (1).

Vidulich, M.A. (2000). Testing the Sensitivity of Situation Awareness Metrics in Interface Evaluations. In: M.R. Endsley & D.J. Garland, *Situation Awareness Analysis and Measurement*. London: Lawrence Erlbaum Associates.

Wickens, C.D., Mavor, A.S. & McGee, P. (Eds) (1997). *Flight to the future. Human factors in air traffic control*. Commission on Behavioral and Social Sciences and Education, National Research Council, National Academy Press.

Page intentionally left blank

ABBREVIATIONS AND ACRONYMS

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

ATC	Air Traffic Control
ATCC	Air Traffic Control Centre
ATCO	Air Traffic Control Officer / Air Traffic Controller *(UK/US)*
ATM	Air Traffic Management
CRM	Crew resource Management
CRNA	Centre Régional de la Navigation Aérienne *(France)*
C-SAS	Cranfield Situation Awareness Scale
DERA	See 'QinetiQ'
DAS	Directorate ATM Strategies *(EUROCONTROL Headquarters, SD)*
DAS/HUM *or just* HUM	Human Factors Management Business Division *(EUROCONTROL Headquarters, SD; formerly known as 'DIS/HUM' or just 'HUM')*
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')*
DST	Decision Support Tool(s)
EATCHIP	European Air Traffic Control Harmonisation and Integration Programme *(now EATM(P))*
EATM(P)	European Air Traffic Management (Programme) *(formerly EATCHIP)*
EEC	EUROCONTROL Experimental Centre *(Brétigny, France)*
FAA	Federal Aviation Administration *(US)*
GUI	Guidelines *(EATCHIP/EATMP)*

HFFG	Human Factors Focus Group *(EATM, HRT; formerly known under EATMP as 'HFSG' standing for 'Human Factors Sub-Group')*
HMI	Human-Machine Interface
HRS	Human Resources Programme *(EATMP)*
HRT	Human Resources Team *(EATCHIP/EATM(P))*
HSP	Human Factors Sub-Programme *(EATMP, HRS)*
HUM	Human Resources (Domain) *(EATCHIP/EATMP)*
IANS	Institute of Air Navigation Services *(EUROCONTROL, Luxembourg)*
ICAO	International Civil Aviation Organization
IFATCA	International Federation of Air Traffic Controllers' Associations
IMPACT (Model)	Integrated Model of Perceived Awareness Control (Model)
MSP	Multi-sector Planning
OJT	On-the-Job Training
OJTI	On-the-Job-Training Instructor
PC	Planning Controller
PCT	Perceptual Control Theory
QinetiQ	*Formerly known as the UK 'Defence Evaluation and Research Agency' (DERA)*
REP	Report *(EATCHIP/EATMP)*
R/T	Radiotelephony
SA	Situation/al Awareness
SA/BARS	Situation Awareness Behaviourally Anchored Rating Scales
SABET	Study of an ATC Baseline for the Evaluation of Team configurations
SAGAT	Situation Awareness Global Assessment Technique

SALIANT	Situation Awareness Linked Indicators Adapted to Novel Tasks
SALSA	Situation Awareness bei Lotsen der Streckenflugkontrolle im kontext von Automatisierung *(situation awareness of en-route air traffic controllers in the context of automation)*
SAPS	Situation Awareness ProbeS
SART	Situational Awareness Rating Technique
SASHA	Situation Awareness for SHAPE
SASHA_L	Situation Awareness for SHAPE on-Line
SASHA_Q	Situation Awareness for SHAPE Questionnaire
SAVANT	Situation Awareness Verification Analysis Tool
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) *(EATMP, HRS, HSP)*
SME	Subject Matter Expert
SPAM	Situation Present Assessment Method
TC	Tactical Controller
TRM	Team Resource Management

Page intentionally left blank

ACKNOWLEDGEMENTS

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.

Many people have helped in the development of the situation awareness measures, particularly the SASHA Questionnaire. The authors gratefully acknowledge assistance provided by:

- Patrick O'Doherty, Björn Royne and John Hyde (EUROCONTROL Institute of Air Navigation Services [IANS], Luxembourg);
- Alain Signamarcheix and other controllers of the Centre Régional de la Navigation Aérienne (CRNA) Sud-Est (Aix-en-Provence ATCC);
- Bert Ruitenberg (International Federation of Air Traffic Controllers' Associations [IFATCA]).

The authors also wish to thank the following people who provided much useful information, either about specific situation awareness measures whose development they have been personally involved with, or about their research:

- Ben Willems (Federal Aviation Administration [FAA] Technical Center, http://www.tc.faa.gov/act-500/hfl);
- Frank Durso (Texas Technical University, http://www.ttu.edu/~psyhfp)[11];
- Yorck Hauß (Technical University of Berlin);
- Michael Vidulich (Armstrong Laboratory, Wright-Patterson Air Force Base, Ohio);
- Alexis Fink (Old Dominion University, Norfolk, VA);
- Christine Boag (University of Queensland).

Finally, our understanding of situation awareness was usefully informed by interviews with controllers at the Athis-Mons ATCC, Philippe Renaud and the members of his team.

Document configuration

Carine Hellinckx *(External contractor)*	EUROCONTROL Headquarters, DAS/HUM

[11] Professor Frank Durso was previously based at the University of Oklahoma.

Page intentionally left blank

APPENDIX A: SASHA QUESTIONNAIRE

Introduction

Computer-assistance tools and other forms of automation support are being increasingly introduced into today's Air Traffic Management (ATM) systems, and are expected to be fundamental components of systems in the future. The success of such automated tool support will depend in part on the degree to which Human Factors are taken into account in the design and implementation of these tools.

As part of the overall European ATM Programme (EATMP) the Human Factors and Manpower Unit[12] within EUROCONTROL has recently initiated a new programme of work to address the human factors issues of automation in ATM systems. The programme is called ‘SHAPE’ (for ‘Solutions for Human-Automation Partnerships in European ATM’). The present aim of SHAPE is to develop a number of measurement techniques that can be applied during real-time simulations to assess and measure the effectiveness of the automation.

This questionnaire is concerned with measuring your ‘situation awareness’. It consists of ten questions. Please answer each question by ticking the box as appropriate. Add any other comments in the space provided.

Thank you for your assistance and cooperation.

SASHA Questionnaire
Q1: - Did you have the feeling that you were ahead of the traffic, able to predict the evolution of the traffic? Never ☐ ☐ ☐ ☐ ☐ Always Comments:…………………………………………………………………………………………… ……………………………………………………………………………………………………… ………………………………………………………………………………………………………
Q2: - Did you have the feeling that you were able to plan and organise your work as you wanted? Never ☐ ☐ ☐ ☐ ☐ Always Comments:…………………………………………………………………………………………… ……………………………………………………………………………………………………… ………………………………………………………………………………………………………
Q3: - Have you been surprised by an a/c call that you were not expecting? Never ☐ ☐ ☐ ☐ ☐ Often Comments:…………………………………………………………………………………………… ……………………………………………………………………………………………………… ………………………………………………………………………………………………………

[12] today known as ‘Human Factors Management Business Division (DAS/HUM)’.

Q4: - Did you have the feeling of starting to focus too much on a single problem and/or area of the sector?

Never ☐ ☐ ☐ ☐ ☐ Often

Comments:……………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………

Q5: - Did you forget to transfer any aircraft?

Never ☐ ☐ ☐ ☐ ☐ Often

Comments:……………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………

Q6: - Did you have any difficulty finding an item of (static) information?

Never ☐ ☐ ☐ ☐ ☐ Always

Comments:……………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………

Q7: - Do you think the <*name of tool*> provided you with useful information?

Never ☐ ☐ ☐ ☐ ☐ Always

Comments:……………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………

Q8: - Were you paying too much attention to the functioning of the <*name of tool*>?

Never ☐ ☐ ☐ ☐ ☐ Always

Comments:……………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………

Q9: - Did the <*name of tool*> help you to have a better understanding of the situation?

Never ☐ ☐ ☐ ☐ ☐ Always

Comments:……………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………

Q10: - Finally, how would you rate your overall situation awareness during this exercise?

Poor ☐ Quite poor ☐ Okay ☐ Quite good ☐ Very good ☐

Comments:……………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………

APPENDIX B: INSTRUCTIONS FOR SASHA ON-LINE

These instructions are intended for the person(s) who will apply the SASHA on-Line situation awareness measure. It is assumed that you are an experienced air traffic controller (either operational or ex-operational). Ideally, you will have some experience of On-the-Job Training (OJT) and therefore be familiar with evaluating another controller's performance.

It is assumed that you understand what is 'situation awareness', what is the purpose of the measure and how it is to be used. (It is assumed that you have briefed beforehand by the EEC Project Team or human factors specialists assigned to the simulation.)

Before the simulation (or measured exercises)

1. Familiarise yourself with the simulation system, in particular any new computer-assistance tools or other forms of automation support that are being evaluated.

2. Ensure that you have read all of the information provided to the participating controllers and attend the pre-simulation briefings. Familiarise yourself with the airspace, route structure, and traffic samples being simulated.

3. Familiarise yourself with the SASHA workstation. (This workstation will either be located in the pseudo-pilots room, in the Human Factors Laboratory, or other facility separate from the main simulation room.)

During the simulation

4. During each specified simulation exercise, as directed by the Project Team leader or other specialist responsible for coordinating the simulation, ask the specified controller your set of situation awareness questions. Each question should be recorded on the SASHA_L query form (see attached) and scored appropriately.

5. When formulating the questions to ask the following general guidelines should be followed:

   - It is important to remember at the outset that the aim of the exercise is to test the other controller's situation awareness from the operational point of view. The aim is not to see if you can force him/her to answer incorrectly.

   - It is important to ask the question at an appropriate moment, both operationally and from the point of view of not disturbing the controller too much.

   - One query should be made about every five minutes (i.e. twelve questions overall during a one-hour exercise). The timing does not need to be precise - an interval varying between three to ten minutes will be satisfactory.

   - The queries should be worded clearly and concisely (not forgetting that English might not be the native language of the controller) such that it is reasonable to expect the controller to give a clear and concise answer (see examples below).

   - Some of the queries - about one-third - should be related to information displayed by the automation tools (e.g. questions 2, 4, 8 and 12 in example given below).

- Some of the queries – about one third – should be related to the evolution of the situation in the near future (e.g. questions 1, 5, 6 and 9 in example given below). How long in advance might be dependent on the potential extrapolation capabilities of the tools provided.
- Some of the queries - again about one third - should be related to knowledge about the current situation (e.g. questions 3, 7, 10 and 11 in example given below).
- Typically the queries will all be different, but a query can be repeated if it is felt to be useful.

6. When scoring the controller's answer:
    - Evaluate the response according to the situation at the moment it is made.
    - Evaluate the operational accuracy of the answer, rather than its absolute correctness, i.e. does his/her answer confirm that the controller has good SA and indicates good ATC?
    - When evaluating the time taken to answer take into account both the complexity of your question and the probable activity/task that the controller was doing when you asked it.
    - A too quick answer might be the sign of a lucky guess. If you feel this is the case then tick the box 'Too Short ☐'.
    - It is important to make your evaluation and scoring as soon as possible after the controller has answered while the exact circumstances are still fresh in the memory.

<u>After the simulation</u>

7. After the exercise, review your completed Query forms and put down any additional comments. It might be useful to discuss some points with the controller concerned, but make sure you get the help and approval of the simulation analyst.

<u>SASHA on-Line Query Form</u>

(photocopy as necessary)

SASHA on-Line Query N°. ……			
Query:……………………………………………………………………………………			
Query's operational importance	- |__|__|__|__|__|__|__|__| +		
Answer's operational accuracy	Incorrect ☐	OK ☐	Correct ☐
Time to answer	Too Short ☐	OK ☐	Too Long ☐

SASHA on-Line Query N°. ……			
Query:……………………………………………………………………………………			
Query's operational importance	- |__|__|__|__|__|__|__|__| +		
Answer's operational accuracy	Incorrect ☐	OK ☐	Correct ☐
Time to answer	Too Short ☐	OK ☐	Too Long ☐

SASHA on-Line Query N°. ……			
Query:……………………………………………………………………………………			
Query's operational importance	- |__|__|__|__|__|__|__|__| +		
Answer's operational accuracy	Incorrect ☐	OK ☐	Correct ☐
Time to answer	Too Short ☐	OK ☐	Too Long ☐

SASHA on-Line - Examples of Queries

No.	Time	Examples of SASHA on-Line queries
1	15:01	Will US Air 1650 and Continental 707 be in conflict if no further action is taken?
2	15:06	Which sector, shown in the communication < *tool* > window has requested a change of FL at handover?
3	15:10	Are there any speed-conflicts on the J74 airway?
4	15:18	What is the time of the situation displayed in < *tool* > window?
5	15:23	Are you expecting any significant increase of workload within the next < *15* > minutes?
6	15:29	Which aircraft needs to be transferred next?
7	15:35	Which aircraft has the faster ground speed, US Air 992 or Air France 2249?
8	15:40	Which of the two conflicts shown in < *tool* > is more critical?
9	15:46	Which aircraft would benefit from a direct route: BA1814 or AF5210?
10	15:51	Which aircraft is going to reach its Requested Flight Level first: AA369 or US Air 551?
11	15:54	With which sector do you need to coordinate AF222 Exit Level?
12	15:59	Which of the two conflicts shown in < *tool* > is more critical?