EUROPEAN ORGANISATION
FOR THE SAFETY OF AIR NAVIGATION

EUROCONTROL

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

Guidelines for Trust in Future ATM Systems: A Literature Review

Edition Number	**:**	**1.0**
Edition Date	**:**	**05.05.2003**
Status	**:**	**Released Issue**
Intended for	**:**	**EATMP Stakeholders**

EUROPEAN AIR TRAFFIC MANAGEMENT PROGRAMME

DOCUMENT CHARACTERISTICS

TITLE

Guidelines for Trust in Future ATM Systems: A Literature Review

	EATMP Infocentre Reference:	030317-01
Document Identifier	**Edition Number:**	1.0
HRS/HSP-005-GUI-01	**Edition Date:**	05.05.2003

Abstract

The purpose of this document and its sequels is to provide a set of human factors guidelines for facilitating and fostering human trust in Air Traffic Management (ATM) systems. The guidelines are primarily concerned with the trust of computer-assistance tools and other forms of automation support, which are expected to be major components of future ATM systems.

This deliverable, on the subject of trust guidelines, is the first one developed within the 'Solutions for Human-Automation Partnerships in European ATM (SHAPE)' Project. There are two subsequent deliverables on trust issues; one is dealing with the measurement of trust (see EATMP, 2003a), the other provides detailed info about trust principles (see EATMP, 2003b).

Keywords

Air Traffic Control (ATC)	Air Traffic Management (ATM) system	Automation
Complacency	Computer assistance	Guidelines
Human Factors	Human-machine	Measure
Rating scale	Solutions for Human-Automation Partnerships in European ATM (SHAPE)	System
Trust	Understandability	

Contact Persons	Tel	Unit
Oliver STRAETER, SHAPE Project Leader	+32 2 7295054	Human Factors & Manpower Unit (DIS/HUM)
Michiel WOLDRING, Manager, HRS Human Factors Sub-Programme (HSP)	+32 2 7293566	Human Factors & Manpower Unit (DIS/HUM)

Authors

C. Kelly, M. Boardman, P. Goillau and E. Jeannot

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	15.05.03
Chairman HRT Human Factors Sub-Group (HFSG)	V.S.M. WOLDRING	16 may 2003
Manager EATMP Human Resources Programme (HRS-PM)	M. BARBARINO	16/05/03
Chairman EATMP Human Resources Team (HRT)	A. SKONIEZKI	16/5/2003
Senior Director Principal EATMP Directorate (SDE)	W. PHILIPP	14.05.03

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	20.12.2000		Working Draft	All
0.2	27.02.2001		First Draft	All
0.3	22.06.2001		Second Draft	1.2, 4.2, 5.4, 6.4, References
0.4	13.02.2002		Approval by HFSG7	All
0.5	30.08.2002		Proposed Issue for HRT18	All (document configuration)
1.0	05.05.2003	030317-01	Released Issue	All (document configuration)

CONTENTS

EXECUTIVE SUMMARY

This document with its sequels provides human factors guidelines for facilitating and fostering human trust in ATM systems. In particular, it is concerned with the trust of computer-assistance tools and other forms of automation support, which are expected to be major components of future ATM systems.

It contributes to the first part of a larger project entitled 'Solutions for Human-Automation Partnerships in European ATM (SHAPE)' being carried out by the ATM Human Resources Unit of EUROCONTROL, later renamed the Human Factors and Manpower Unit (DIS/HUM).

The former UK Defence Evaluation and Research Agency (DERA), now known as QinetiQ, was awarded the investigation of three specific human factors topics concerned with trust (see the current document and EATMP, 2003a, 2003b), situation awareness (see EATMP, 2003c) and teamworking (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).

This deliverable, on the subject of trust guidelines, is the first one for the SHAPE Project. There are two subsequent deliverables on trust issues; one is dealing with trust measures (see EATMP, 2003a), the other provides detailed information about trust principles (see EATMP, 2003b).

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

Section 2, 'Automation', defines what is meant by automation, introduces the concept of different 'levels' of automation and provides a brief review of human factors research.

Section 3, 'Understanding Trust', defines what is meant by trust, discusses the different elements or dimensions which trust is composed of, and explains the notion of 'complacency'. How trust is understood in ATC is discussed.

Section 4, 'Trust and Human-Machine Systems', provides a short review of human factors research into trust and automation. The first part considers general research; the second part focuses on research into trust and ATM systems. Lastly, a simple model of trust factors is presented.

Section 5, 'Measuring Trust', describes a number of techniques for measuring trust.

Section 6, 'Trust Subjects', provides subjects to be considered for facilitating and promoting trust in the design and development of ATM systems.

References, a list of the Abbreviations and Acronyms used in these guidelines and their full designations, and finally Acknowledgements are provided at annex.

Page intentionally left blank

1. INTRODUCTION

1.1 Purpose

The purpose of this document and its sequels is to provide a set of human factors guidelines for facilitating and fostering human trust in ATM systems. The guidelines are primarily concerned with the trust of computer-assistance tools and other forms of automation support, which are expected to be major components of future ATM systems.

ATM systems have always depended on trust. Controllers have to trust their radar and communications equipment, trust the safety of their procedures and instructions and, ultimately, trust pilots and others to follow those instructions correctly. However, with the introduction of computer assistance 'tools' and other forms of automation support, trust is becoming more important because **it is potentially both harder to gain and easier to lose.**

1.2 Scope

This document is intended to provide a review of the literature on the subject of human trust in automation, particularly in relation to the real-time simulation of future ATM systems. The review is not intended to be exhaustive. In addition, the human factors guidelines are aimed at providing practical advice to EUROCONTROL project leaders and other project staff who are concerned with ATM automation design issues.

Trust is recognised to be a subject of increasing importance in the design of complex human-machine systems. However, trust is not a simple uni-dimensional variable. It is possible to be correctly distrusting of a system (e.g. when it is unreliable), but also to be too trusting (‘over-trusting’) or not trusting enough (‘under-trusting’). These four different aspects of trust are illustrated in Figure 1.

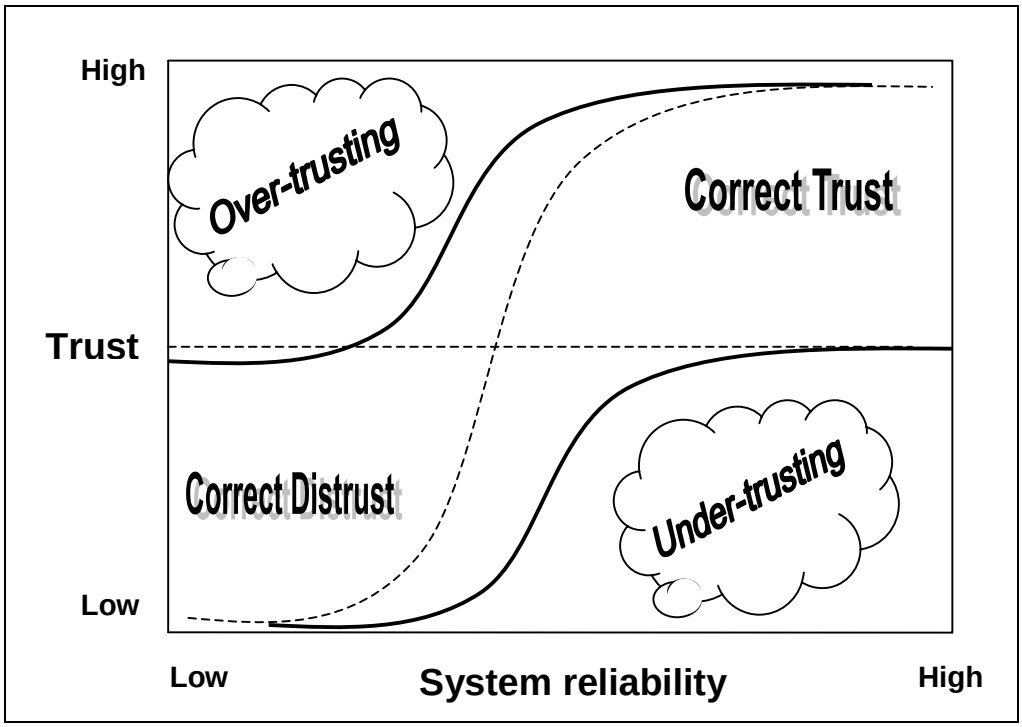

Figure 1: Theoretical relationship between trust and system reliability

Trust is affected by reliability, but is clearly not the same since the latter is a property of the automation and the former a property of the human. Trust is important because, ultimately, we want controllers to use automation support when the latter are useful and reliable.

1.3 Background

The work on trust presented 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 the Human Factors and Manpower Unit (DIS/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 this document and EATMP, 2003a, 2003b).

- 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 (see EATMP, 2003c).

- 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). 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 six sections, following the 'Introduction', as shown in Figure 2.

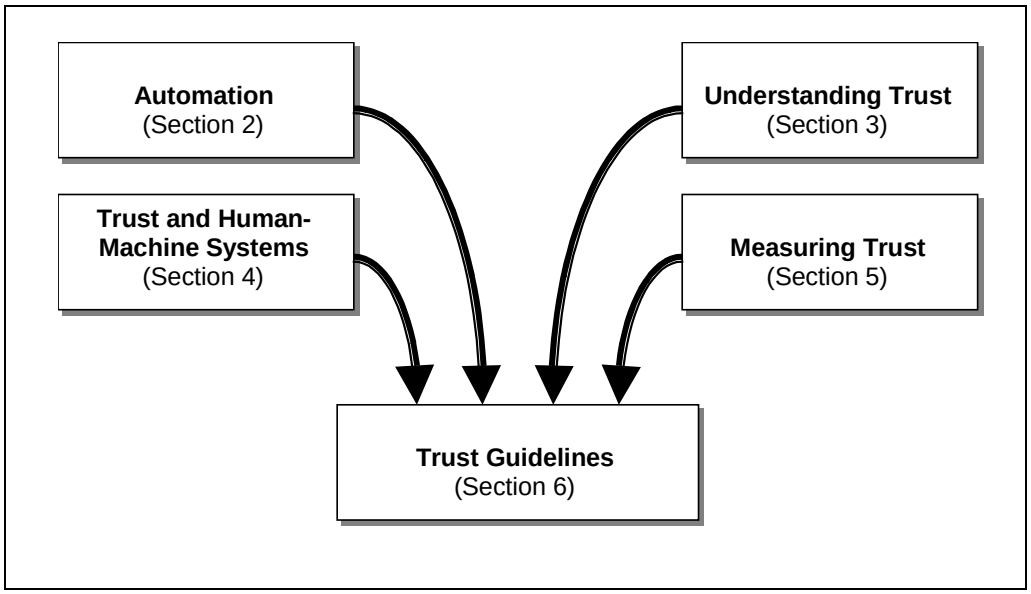

Figure 2: Structure of the guideline document

2. AUTOMATION

2.1 What is Automation?

In order to understand the concept of trust in automation, it is important first to understand what is meant by the term *automation*. A variety of meanings can be found in the literature. At one extreme, automation is often equated with technological change in general, particularly changes that result in humans being replaced by machines. At the other extreme automation is associated with computers and computer software, particularly including 'intelligent' systems that are capable of a degree of self-operating (autonomous) functioning.

In the context of human-machine systems, and especially ATM systems, a more precise definition of automation is required. A good definition, which will be used here, has been provided by the US National Research Council, Panel on Human Factors in Air Traffic Control Automation[1]:

> *… a device or system that accomplishes (partially or fully) a function that was previously carried out (partially or fully) by a human operator. (p. 243).*

This definition emphasises a number of important points:

- First, automation should be distinguished from technological innovation or modernisation, which, although involving computerisation, does not necessarily change the *allocation of functions* between humans and machines. For example, replacing a radar display with a high-resolution, colour, computer display terminal is not automation if the controller's tasks remain unchanged. However, according to this definition, replacing paper flight progress strips with an electronic equivalent is automation because the allocation of functions (e.g. coordination between sectors) has changed.

- Second, automation is not a single entity or attribute of a system. The word 'partially' in the above definition indicates that automation can be applied to different degrees or levels. For example, the autopilot and auto-landing functions in modern aircraft represent one level of automation. In the case of electronic flight progress strips, the automatic sorting of the strips on the display screen would be another level of automation. The subject of levels of automation is considered in more detail in the next section.

- Third, as noted by Parasuraman and Riley (1997), what is considered to be automation changes over time as technology changes.

[1] See Committee on Human Factors Web site at:
http://www4.nas.edu/cbsse/delhp.nsf/web/committee_on_human_factors

- Ultimately, if a function is fully accomplished by a device or system, then the function ceases to be regarded as automation, and is instead seen as an ordinary part of system. In other words “today’s automation could well be tomorrow's machine” (p. 231).

2.2 Levels of Automation

The point was made above that automation could be applied at different levels. It is useful to think of automation as a continuum varying from manual control at one end, to fully automatic at the other end. This notion of levels has been discussed in the literature for many years (e.g. Hopkin, 1975). The ten levels of automation first expounded by Thomas Sheridan (e.g. Sheridan, 1988) have become widely accepted. The most recent version of these levels, from Parasuraman, Sheridan and Wickens (2000), is shown in Table 1 below.

Table 1: Levels of automation (from Parasuraman *et al.*, 2000)

Level	Human-machine cooperation
10 (high)	The computer decides everything, acts autonomously, ignoring the human.
9	Informs the human only if it, the computer, decides to.
8	Informs the human only if asked, or
7	Executes automatically, then necessarily informs the human, and
6	Allows the human a restricted time to veto before automatic execution, or
5	Executes that suggestion if the human approves, or
4	Suggests one alternative
3	Narrows the selection down to a few, or
2	The computer offers a complete set of decision/action alternatives, or
1 (low)	The computer offers no assistance: human must take all decisions and actions.

It is interesting to note that even with full automation (level 10) the need for (human) maintenance and modification of the system is unlikely to disappear. Indeed, as discussed by Shorrock and Scaife (2000), today’s modern Air Traffic Control Centre (ATCC) has necessarily a complimentary control and monitoring system to monitor its functioning. Engineers must continuously operate this monitoring system, not unlike the control room of a nuclear power station. Therefore, as discussed by Bainbridge (1982), one can draw the paradoxical conclusion that an automated system is still a human-machine system.

2.3 Human Factors Research

The subject of automation has attracted, justifiably, much research over the past two decades. Much of this research has been devoted to aviation issues and particularly automation of aircraft cockpits and the flight decks of commercial airlines (Wiener & Curry, 1980; Wiener, 1985; Billings, 1991; ICAO, 1994; Funk *et al.*, 1996). However, as indicated by two recent reviews of current research (Mouloua & Koonce, 1997; Scerbo & Mouloua, 1999), it is very evident that automation now permeates all human-machine systems from driving a motor car, to anaesthesiology. Automation raises a host of human factors issues, not only trust (the subject of this guideline document), but also error, situation awareness, vigilance, stress and workload. A theoretical analysis of the human use of automation, in particular its *misuse*, *disuse* and *abuse*, has been put forward by Parasuraman and Riley (*op. cit.*). They argue that system design can be improved by a proper understanding of the human factors associated with each of these aspects.

The automation of ATM systems, and particularly ATC systems, has also received significant attention over several decades (e.g. Hopkin, 1975; Erzberger, 1989; Wise *et al.*, 1991; Hollnagel *et al.*, 1994), though perhaps less than for other domains. This is not so surprising when one considers that despite undergoing many developments over past decades, ATM systems remain remarkably *unchanged*. That is, the controller's basic 'tools' of radar, flight progress strips and Radiotelephony (R/T) remain the cornerstones of ATCCs throughout the world.

For the FAA, Cardosi and Murphy (1995) conducted a major review of automation issues in ATC, discussing both its benefits and disadvantages and including a checklist of design recommendations. The US National Research Council, Committee on Human Factors (see also 2.1 above) carried out a wide-ranging study of the human factors issues of ATC systems and technology, focusing particularly on automation (Wickens *et al.*, 1997). The impetus for the study was the concern that

> *efforts to modernize and further automate the air traffic control system should not compromise safety by marginalizing the human controller's ability to effectively monitor the process, intervene as spot failures in the software or environmental disturbances require, or assume manual control if the automation becomes untrustworthy.* (p. ix)

The effects of automation on the future role of the controller is also a topic of much research (e.g. Kelly & Goillau, 1996; Hopkin, 1998; Cox & Kirwan, 1999). It was a particular concern of the 'Programme for Harmonised Air Traffic Management Research in EUROCONTROL (PHARE)' (EUROCONTROL, 1993) and of the European Commission's 'Role of the Human in the Evolution of ATM Systems (RHEA)' Project (Nijhuis, 1998). It is beyond the scope of this document to discuss automation in greater detail, but future controller roles and other issues of automation will be the focus of SHAPE work on skill set requirements.

Automation is a device or system that accomplishes (partially or fully) a function that was previously carried out (partially or fully) by a human operator.

3. UNDERSTANDING TRUST

3.1 What is Trust?

Trust is a term that is familiar to all of us in everyday life. We talk of the trust that we have other people (family, friends and colleagues), how much we believe what we see or are told (e.g. in a newspaper), or how confident we are that something works properly (e.g. a motor car). Clearly, trust has several meanings, but it can be defined simply as the confidence placed in a person or thing, or more precisely, the degree of belief in the strength, ability, truth or reliability of a person or thing. In the context of complex, human-machine systems, Madsen and Gregor (2000) have defined trust as follows:

> *Trust is the extent to which a user is confident in, and willing to act on the basis of, the recommendations, actions, and decisions of an artificially intelligent decision aid.*

This is a useful definition for the purposes of SHAPE. However, the terms 'artificially intelligent' suggest too strongly that the focus is upon expert systems and related computer systems. Therefore, a term such as computer-based tool is preferred.

In psychological jargon trust is an *intervening* variable because it 'intervenes' between particular stimulus conditions and particular behaviours. That is, it is an internal state that cannot be measured directly, but is inferred on the basis of certain observations and measurements. In the context of SHAPE, the degree of trust in automation could, theoretically at least, be inferred from objective measures of controller performance (e.g. frequency, accuracy or speed of interaction), if the relationship between these measures and the automation could be unequivocally established.

An intervening variable such as trust can also be measured *subjectively* by asking an operator or controller to say simply how they feel. Indeed, as described below (see 5.2), the use of subjective rating scales is the most common means of measuring trust. It is important to note that if the origin of the subjective ratings can be modelled, one can convert intervening variables to objective measures. The potential development of an objective measure of trust is discussed later (see 5.3).

Trust is the extent to which a user is willing to act on the basis of, the recommendations, actions, and decisions of a computer-based 'tool' or decision aid.

3.2 Dimensions of Trust

In order to provide guidance about how best to take account of trust in the design of human-machine systems, it is essential to understand the nature of trust. From a psychological perspective, trust is a construct that is composed of several different elements or dimensions; these dimensions need to be examined. It is also important to understand the relationship of trust to the two other measures proposed for SHAPE, i.e. situation awareness and teamworking. Clearly, these three measures are not independent of each other.

As indicated above, trust is a subject that pervades many aspects of daily life, particularly inter-personal relationships. Three broad areas of research that have explored the dimensions of trust can usefully be distinguished. These are:

1. Social psychology (i.e. human-to-human interaction).
2. Systems engineering (i.e. human-machine interaction).
3. Information technology (i.e. machine-mediated, human-human interaction).

Social psychology

A model of trust in close relationships was developed by Rempel, Holmes and Zanna (1985) consisting of three components: **predictability**, **dependability** and **faith**. There are two interesting aspects to this model. First, whereas predictability and dependability are related to past experience and reliability of previous evidence, the third component of faith is concerned with generalising to future situations (i.e. going beyond the available evidence). Second, the model is hierarchical. The first stage of trust (predictability) focuses upon concrete, specific behaviours; the second stage (dependability) is concerned with the qualities and characteristics attributed to the other person; and lastly, in stage three (faith), the focus is not upon specific behaviours but beliefs and convictions about future events.

Trust in organisations has also been extensively studied (e.g. Kramer, 1999).

Systems engineering

By systems engineering is meant human-machine systems in the broadest sense of that term, i.e. supervisory and process control systems, command and control systems, air traffic control systems, and so on.

Sheridan (1988) proposed seven attributes or causes of trust that, he claimed, should enable trust to be defined operationally, measured, and modelled. As shown in discussion of measures of trust (see Section 5), this claim has largely been borne out. The seven attributes are: **reliability**, **robustness**, **familiarity**, **understandability**, **explication of intention**, **usefulness** and **dependence**.

The experimental investigations by Muir (1994) and Muir and Moray (1996) showed that the Rempel *et al.* Model of trust between humans could usefully be applied to the development of trust between humans and machines. In addition, Muir found evidence that the expectation of machine **competence** (i.e. the extent to which it does its job properly), best captured what operators meant by saying that they trusted a machine.

Lee and Moray (1992, 1994) developed Muir's studies further. In particular the authors showed that operators' reliance on automation depends not only on trust in the automation, but also on the operators' **self-confidence** in their own abilities.

Information technology

In contrast to the research on trust within the context of complex, safety-critical, semi-automated systems, another research area has looked at human trust of computer systems within the context of the information technological society. For example, Hall (1996) discusses the notion of an 'assistant-like interface' for computers which users can trust to behave in accordance with their goals and priorities. Abdul-Rahman and Hailes (1999, 2000) have proposed a sociologically based model of trust for the problem of reliable information retrieval. Their model incorporates the notion of **reputation** that is then generalised so that reputational information can come from either an external source, or from the truster himself through experiences with other 'agents'.

Information technology offers the potential of giving people the opportunities to communicate in a multitude of different ways in different social and business networks. However, fulfilling this promise relies on the use of computers, and most critically it depends on the design of, what has been called in a recent National Research Council study (NRC, 1997), 'Every-Citizen Interfaces' (ECIs). According to the latter, at least four different facets of trust arise when considering collaboration and communications in ECIs:

1. Privacy - trusting the system with information. This concerns not only privacy for the individuals using systems, but also for developers.

2. Authentication - trusting what the system reports about users.

3. Credibility - trusting the content of the system.

4. Reliability - trusting the system to function.

The trustworthiness of networked information systems is a huge and growing topic of research. Although much of this research is focused upon software reliability, availability, privacy and security, it does touch upon some relevant human factors issues. In addition, the intriguing idea of building 'trustworthy systems from untrustworthy components' (Schneider, 1999) is a research topic worth of further study.

Trust is a construct composed of several elements or dimensions. The main dimensions identified in the research literature are:

- ***Predictability***
- ***Dependability***
- ***Faith***
- ***Reliability***
- ***Robustness***
- ***Familiarity***
- ***Understandability***
- ***Explication of intention***
- ***Usefulness***
- ***Competence***
- ***Self-confidence***
- ***Reputation***

3.3 Complacency

Mention will be made about the notion of *complacency* by which is meant that operators of a highly reliable system, such as aircraft flight deck, will fail to monitor it sufficiently so as to detect faults, and will become 'complacent'. The empirical work most often cited in respect of this, is that of Parasuraman *et al.* (1993). More recently, Parasuraman and Riley (1997) have referred to over-reliance on automation as an example of automation 'misuse'.

Whilst it is true that both experiments and field studies show evidence of operators missing signals, it is not clear that this evidence supports the notion of complacency. Indeed, a re-analysis by Moray (1999) has shown that, on the contrary, the operators used by Parasuraman *et al.* (*op. cit.*) were most probably behaving in an optimal manner and therefore did not *detect* some of the faults. Because the research focused on the detection of signals, not on sampling, it only appears that the behaviour was complacent. However, one cannot measure complacency by detection. Moreover, it can be shown that even optimal sampling cannot ensure the detection of all faults or other important signals. It is not complacent, but rational to reduce the frequency of monitoring when observing a highly reliable automated system.

The implications of this analysis for system design are important. According to Moray (*op. cit.*) they imply that in the design of an advanced, highly complex system such as automated ATC, appropriate attention *must* be given to alarms, warnings, and display alerts. Of course, such alarms must be more reliable than the system whose failures they signal, and at the same time they must not produce many false alarms or they will not be trusted. In addition, attention must be given to training, but training can only guarantee optimal monitoring, not safety.

Three kinds of monitoring behaviour can usefully be distinguished (Moray, Inagaki & Parasuraman, 2001). If operators sample a variable *less* often than is demanded by an optimal attention strategy it can be called 'complacent'; if operators sample *more* often than is demanded it may be called 'sceptical'.

What is needed is a balance between complacency and scepticism, that is, between under-sampling and over-sampling. An accurately trained operator who is trained to have exactly the correct sampling behaviour may be called 'eutactic'.

However, as noted by Moray, Inagaki and Parasuraman (*op. cit.*), when dealing with real systems such as air traffic control, power stations, aircraft, motor cars, etc., we are not content with eutactic behaviour, let alone complacent behaviour. What we intuitively require is rational sceptical behaviour – but at what frequency?

'Complacency' is a term used to describe an operator's over-reliance on automation resulting in the failure to detect system faults or errors. Complacency is also referred to as one kind of automation 'mis-use'.

3.4 Trust and ATC

A fundamental premise of the SHAPE Project is that the concept of trust and the dimensions of trust (as described above) are equally applicable to the domain of ATC as they are for other domains such as industrial process control. Is this assumption really valid? At first sight it seems intuitively obvious that, given the safety-critical nature of ATC, trust is an intrinsic part of the controller's job. As Hopkin (1995) put it succinctly, *air traffic control depends on trust … Pilots have to trust controllers to issue instructions that are safe and efficient. Controllers have to trust pilots to implement those instructions correctly. Both have to trust their equipment, their information sources and displays, their communications, and the safety of their procedures and instructions* (p. 346).

On the other hand this view of trust and ATC does not entirely accord with the controllers' perception of the subject. As discussed later (see 4.2), there is some evidence that controllers do not think in terms of 'trusting' the system that they work with. Instead, their concern is more with the operational reliability of the system (and its tools) which either works or does not. If it fails in a critical manner, it will not be used (trusted) again.

A perfect example of this operational viewpoint is provided in connection with the Short-term Conflict Alert (STCA) that has been operational in several ATCCs throughout Europe since the mid-nineties. In the UK extensive trials and tests of the STCA were carried out before it went into operational service (in March 1996). Since then, the performance of the STCA is regularly evaluated, including detailed questionnaire surveys of controllers' opinions. In none of the reports about STCA (e.g. Hale & Baker, 1990; Du Boulay *et al.*, 1994) does the word 'trust' appear! Instead, the focus is upon the operational

acceptability of STCA (particularly minimising 'nuisance' alerts[2]), its effectiveness (as a safety net), the training implications, and possible display enhancements.

The attitudes of controllers to trust, and particularly their trust of automation, has been the subject of a recent survey conducted by the National Aerospace Laboratory (NLR) of The Netherlands. The results of the survey (EUROCONTROL, 2000a, 2000b) showed controllers were in general positive about automation. For example, over 60% of controllers reported not being mistrusting of technology (i.e. over 60% *disagreed* with the statement that "I do not trust new ATC technology, even though it is designed to make my job easier"). Interestingly, some marked differences were found in the attitudes of management staff compared to the controllers. Management tended to feel that controllers would be distrusting of new technology, such as computerised decision aiding systems, would be confused by extraneous system features, and unable to learn new technology. However, these views simply do not fit with controllers' own self-reports.

The survey confirmed that reliability was highly valued in new systems as a key determinant of controller trust. However, it was noted that 'lower' forms of ATC automation like STCA, despite having known reliability problems had come to be accepted and relied upon. This apparent contradiction provides an important lesson about trust in automation. **So long as controllers understand the limitations of the automation, and can clearly see the benefits of using it, they will trust it.**

Trust is an intrinsic part of air traffic control. Controllers must trust their equipment and trust pilots to implement the instructions they are given. The reliability of new systems is a key determinant of controller trust.

[2] A nuisance alert is one which, although technically correct, is of no practical use and a potential distraction to the controller, because appropriate action has already taken place or the situation is otherwise under control. A nuisance alert is not the same as a 'false alarm' that refers to an alert for no apparent reason, or a 'miss' that refers to the absence of an alert when one should have been triggered.

4. TRUST AND HUMAN-MACHINE SYSTEMS

4.1 General Research

From a review of the literature on human trust in automation two general observations can be made. First, most of the research has been to do with the supervisory control of simulated industrial process systems. This also includes some medical procedures such as anaesthesiology that can reasonably be described as a form of continuous process control (Weinger, 1997). Little of the research has specifically focused upon ATM systems, although a few studies have been found (which are discussed in 4.2). Second, most of the research has investigated trust in the context of how it is affected by faults in the automation.

Empirical research on the subject of trust and human-machine systems can be fairly said to have started with research by Moray's group at Toronto, more particularly with the doctoral studies of Muir (1987, 1994) who investigated trust in the operation of supervisory control systems. Lee and Moray (1992, 1994), Muir and Moray (1996) and, more recently, Moray, Inagaki and Itoh (2000) have extended Muir's pioneering work over the last decade. Amongst the many findings of this extensive work it has been found that trust is strongly affected by system reliability, but self-confidence may or may not be so affected depending on how easily the operators can distinguish their effect on system performance from that of automation.

It has also been shown that it is possible to develop an empirical model of trust based on time series modelling which, as a result, can predict in real time. For example, Lee and Moray (*op. cit.*) were able to predict the probability that (supervisory control) operators would intervene and take over manual control from the automation. Their equation was highly predictive, accounting for over 80% of the variance in some cases. Lee's model also shows that only the recent past affects the level of trust. Effectively, the model says that the operators bring a certain level of trust to the task each time they commence, and then, during work, only the one or two most recent experiences have any effect on the level of trust and self-confidence. These predictive engineering models show that the causal factors driving the dynamics of trust are different from those driving self-confidence. Trust seems to be reduced by *properties* of the system (real or apparent false diagnoses), whereas self-confidence is reduced by experiences of the operator (experiences of accidents).

In an extension of Lee and Moray's study Lewandowsky and colleagues (Tan & Lewandowsky, 1996; Lewandowsky, Mundy & Tan, 2000) compared trust in automation with trust in human partners in equivalent situations. Specifically, participants were required to operate a process control simulation in which some subsystems could be delegated to 'auxiliary' control (either automation or other operators). The key findings were that faults in the automation condition strongly affected trust and self-confidence. Faults reduced trust and subsequent fault-free performance restored it; a similar result was observed

for self-confidence. It was argued that the apparent abruptness, with which trust declined when faults occurred, was intuitively logical in the context of industrial plant operations. Moreover, such a decline in trust could be essential under certain safety-critical situations (e.g. flight deck automation fault), when instantaneous remedial action might be required. The authors concluded *the observed abruptness of trust decline must be borne in mind during the design of automated systems* (Lewandowsky *et al.*, 2000, p.122).

The latter conclusion about the abrupt decline of trust raises an important issue. In certain safety-critical situations, when the system response time or the event time scale is very fast, there may be no time for the operator to respond by re-setting his trust or making decisions. As discussed extensively by Inagaki (e.g. 1999), the implication of this is that in such circumstances the automation must have the over-riding authority to make a decision or act.

As part of Muir and Moray's (1996) study the question of the spread of **distrust** within a system was considered due to its considerable practical importance in the design of automated systems. Muir and Moray wanted to find out whether distrust could spread between components of a system, if so, to which ones, structurally, functionally or causally related, or spread indiscriminately through the entire system. The results of the study showed that distrust could spread between two separate functions of a common physical component. In the study conducted, distrust in a poorly performing pump display affected the levels of trust in the same pump's competent control system. It was also found that distrust in a particular function might spread to other functions performed by the same subsystem. This may lead to unwarranted distrust, unnecessary monitoring and overriding of good decisions. However, distrust did not spread across independent but similar systems, i.e. participants could discriminate between systems, conditioning trust on the particular properties of individual systems.

In a series of experimental studies concerned with combat identification systems, Dzindolet and colleagues (Dzindolet *et al.*, 1999, 2000a) found that providing operators (university students) with information about the conditions in which an automated aid is likely to make errors leads to improved task performance. In addition, depending on the precise experimental condition biases toward (misuse) and against (disuse) automation were observed. The conclusion was drawn that if designers want to encourage human operators to rely on automated systems, they should ensure that operators understand when the aids are likely to make an error. The better this is understood, the more likely that the operator will trust the automation appropriately. The results have been interpreted within a theoretical 'Framework of Automation Use' (Dzindolet *et al.*, 2000b). According to this framework, relative trust (and automation use) is determined from the outcome of a comparison process between the perceived reliability of the automated aid (trust in aid) and the perceived reliability of manual control (trust in self). The outcome of the decision process, termed the perceived utility of the automated aid, will be most accurate when the *actual* abilities are compared. In practice, because the real abilities are not accurately known, errors and biases are likely to occur resulting in the inappropriate (disuse and misuse) of automation. Additional

evidence of the importance of the perceived reliability of automation has been provided by Lui and Hwang (2000).

The importance of feedback to an operator about automation errors was also shown to be the case in a study conducted by Simpson (1992, 1995) in the domain of naval command and control systems. It was found that trust and acceptance was influenced by the provision of explanatory features detailing the underlying decision-model employed by the system and the manner in which it dealt with uncertainty. The operators had neither blind trust nor questioned every decision that the system made, reinforcing Muir and Moray's (1996) conclusions that trust is not a discrete variable but that variable levels of trust can exist between none and total. The explanation facilities assisted the operators in predicting when the system would and would not be correct, thus allowing them to calibrate their trust to specific situations.

The results showed that accuracy and predictability were the most important factors influencing trust in the system. Predictability is influenced by the user's comprehension of the system, which in turn is affected by the decision-making strategies used by the system and the operator. In order for the system to be trusted it must demonstrate technically competent role performance and provide operators with the facilities to enable them to predict the pattern of its accuracy.

Moffa and Stokes (1997), who were investigating operator trust in a medical decision support system, have reported some interesting observations about the effects of the size of errors. Whereas Muir (*op. cit.*) had found that trust decreased as the magnitude of the faults grew (along what looks like an exponential curve), Moffa and Stokes found that only large discrepancies caused medical staff to doubt the system; small discrepancies did not. It was speculated that this was due to high levels of trust and deference to the 'expert' system. Thus, according to Moffa and Stokes (*op. cit.*), relatively minor system errors could, if not corrected or compensated for, lead to a subtle and progressive deterioration in the quality of diagnostic support and guidance. Moffa and Stokes also found evidence to support the hypothesis that the uniqueness of medical diagnoses hinders the development of compensatory strategies when diagnostic error occurs frequently in an expert system (in contrast to an industrial process control context where consistent error may permit such compensating judgements).

Jian *et al.* (1998, 2000) have carried out a series of experimental studies to test the assumption that trust between humans (e.g. Rempel *et al.*, 1985) could be applied to trust between humans and automated systems. (The work of Muir and others mentioned above made this assumption, but it had not been tested empirically.) The results of an extensive questionnaire study and cluster analysis showed that the patterns of ratings were similar across three types of trust: general trust, human-human trust and human-machine trust; that is, for each of the three types of trust, the sets of words related to trust were very similar (see Table 2).

Table 2: Words most related to trust across three trust situations (from Jian *et al.*, 1998)

Conditions		General trust	Trust between people	Trust between human and automated systems
Words	1.	Trustworthy	Trustworthy	Trustworthy
	2.	Honesty	Honesty	
	3.	Loyalty	Loyalty	Loyalty
	4.	Reliability	Reliability	Reliability
	5.	Honour		Honour
	6.		Integrity	
	7.			Familiarity

The results of Jian *et al.* (*op. cit.*) also provided the first empirical evidence that the concepts of trust and distrust could be treated as opposite ends of a trust continuum. In practical terms this implies that trust and distrust can be measured using the same rating scale. In fact, as a result of these experimental studies a multi-dimensional trust scale was developed (see 5.2).

4.2 Trust and ATM Systems

Research on trust in ATM systems is surprisingly limited given its undoubted importance. On the other hand it is evident that the main research focus has been on controller workload and ways to increase traffic capacity without adversely affecting workload. The rationale underlying much recent research has been that the use of computer assistance tools (automation) should enable controller's workload per aircraft to be reduced, thereby releasing 'spare capacity' to increase traffic capacity per controller (Stoner, 1995).

A review of several simulation trials, 'Computer Assistance for En-Route ATC (CAER)', 'Operational Display and Input Development (ODID)' and the PHARE Demonstrations (Kelly *et al.*, 1995; Graham *et al.*, 1994; Whitaker & Marsh, 1997; Reichmuth *et al.*, 1998; Chabrol *et al.*, 1999), shows that the main subjective performance measurements taken were invariably of controller workload, e.g. the NASA Task Load Index (TLX) and Instantaneous Self-Assessment (ISA). No measures of trust were made. However, during the course of these simulation trials controllers' comments on the reliability of various computer assistance tools were noted. These comments provide interesting insights into the problems of trust and automation.

In the PHARE Demonstration 1 (PD/1) trial (Whitaker & Marsh, *op. cit.)* it was noted that controllers, whilst expressing their general approval of the PD/1 operational concept (i.e. of advanced planning), had reservations about the tools. For example, tactical controllers would not always trust the plan for an aircraft trajectory even when generated by their team colleague! The 'Highly Interactive Problem Solver (HIPS)' planning tool could sometimes lead the

planning controllers to plan an aircraft to climb safely in front of another. The tactical controllers would not trust such a plan and monitored these cases closely, expecting loss of separation and thereby possibly increasing their cognitive workload. Mistrust by the controllers of the reliability and accuracy of the tools, was cited as one of the reasons why the expected capacity gains were not demonstrated. It was recommended that this problem (of mistrust) would need to be addressed so as to enable the further development of the PHARE concept and systems.

Similar observations concerning trust were made in the PHARE PD/3 trial (Chabrol *et al.*, *op. cit.*). Although in an automated system such as PD/3 the TC does not need (theoretically at least) to maintain the same mental picture, the importance of the system providing the controller with "trusted and complete trajectory information" on all aircraft trajectories in the sector was highlighted. The information is a fundamental part of cooperation (and coordination) between the controllers. The results from the PD/3 simulation showed that where this information is reduced or breaks down, the automation becomes a hindrance. Furthermore, the fact that the system did not always give conflict information to the controllers or provided sometimes ambiguous and contradictory conflict information induced what was termed 'parasite' behaviour. It reduced controllers' trust in the system and they tended to revert to behaviour they would use without the automation support (behaviour that now became a hindrance). This demonstrates that automated systems can only work if they achieve controllers' trust and anything that detracts from that trust can and will increase workload and decrease safety. It was recommended that for future ATM projects utilising (part of) the PHARE concepts automation should be limited *to what can be without any doubt trusted by controller and provide cooperative system functions allowing the controller to remain the master of the system* (p. 6).

In a small-scale ATC simulation study, Masalonis *et al.* (1998) showed that the consequences of an error might be as important as its size in effecting user trust. The study showed that when an automated aid failed to alert controllers of an impending conflict, the subjective measure of trust was lower than when the aid gave a false alarm. It was concluded that this effect was due to the higher consequence of a miss. This is backed up by earlier research reported by Taylor (1988) considering technical decision-making in fighter aircraft (the 'human-electronic crew'). Subjective ratings showed that the demand for trust was associated with the perceived risk and the probability of negative consequences. Thus relying on another person or system to make risky decisions calls for a large amount of trust. In another study Riley (1994) found that subjects took longer to re-engage automation in high-risk situations than in low-risk situations after a failure of the automated system.

In contrast to the above real-time ATC simulations that have addressed the subject of trust, a different approach was used in a recent European Commission project called 'Role of the Human in the Evolution of ATM Systems (RHEA)'[3]. The general aim of RHEA was to analyse and evaluate a range of 'automation concepts' for guiding the introduction of automation in

[3] See the European Commission web site http://www.cordis.lu/transport/src/rhea.htm.

future ATM systems. Excluding 'full automation' and 'HMI enhancement', seven concepts were distinguished as shown in Table 3.

Table 3: Automation concepts in RHEA Project

Automation concept (from RHEA Project)	Description of automation concept	Equivalent automation level (Table 1)	Automation project
Machine Proposal (MP)	The system proposes options so as to meet high-level system goals (i.e. solutions), which the controller can accept or reject.	2	CORA1 MTCD
Machine-aided Evaluation (ME)	The controller proposes solutions and evaluates them with help of system (e.g. using a 'what-if' tool).	3	
Cognitive Tools (CT)	The controller carries out tasks, but is helped by the system's sophisticated, problem-solving ('cognitive') tools	3-4	CORA2
Dynamic allocation with Human delegation (DH)	Tasks may be done, at different times, either by the controller or system. The *controller* decides when, and what task one or the other will do.	4-5	
Dynamic Aircraft delegation (DA)	Some tasks, e.g. tactical conflict resolution, are delegated from the ground side (controller) to the airborne side (pilot)	4-5	FREER
Dynamic allocation with Machine delegation (DM)	Tasks may be done, at different times, either by the controller or by system. The *system* decides when, and what task one or the other will do.	5-6	
Controller as supervisor	The system performs all tasks. The controller monitors system operation and intervenes in emergencies only.	7-9	CORA3 ?

Each of the automation concepts was evaluated using several different techniques, such as fast-time simulation, human reliability analysis and a modified cognitive walkthrough (Goillau *et al.*, 1998). The latter evaluation, which involved in-depth interviews with four controllers, provided interesting data about the automation concepts, and particularly about trust. The controllers' were very concerned about the trustworthiness of the tools (implicit in the automation concept). Typical of the controllers' comments was this statement: *If the system fails manual reversion is required which could increase stress levels partly due to having to step in and also wondering why the tool has failed. The tool would never be trustworthy again if it failed.* (DERA, 1997, p. 36). The final RHEA report (Nijhuis *et al.*, 1999) included several recommendations that trust in automation tools must be addressed during their design and development.

Lastly, a socio-technical study of trust in ATM systems is currently being conducted at the Applied Psychology Research Group[4] (Trinity College, Dublin) with support from EUROCONTROL. Initial results from a questionnaire and interview survey indicate that controllers view trust in terms of a belief. Moreover, this belief appears to be calibrated quite differently to various system 'referents': self, other people (e.g. controllers, pilots) and technology (Bonini, Jackson & McDonald, 2001).

4.3 A Simple Model of Trust

In the preceding sections of the document various dimensions of trust, and factors influencing trust, have been described. It is useful, particularly when considering how to measure trust (see Section 5), to understand how these various factors are related to each other. A model, or influence diagram, of trust factors is shown in Figure 3.

The model is simple, but allows trade-offs to be made between the different factors (as seems to apply to trust). As well as illustrating the relationship between the trust factors, the model could provide a framework for the resultant design principles and guidelines.

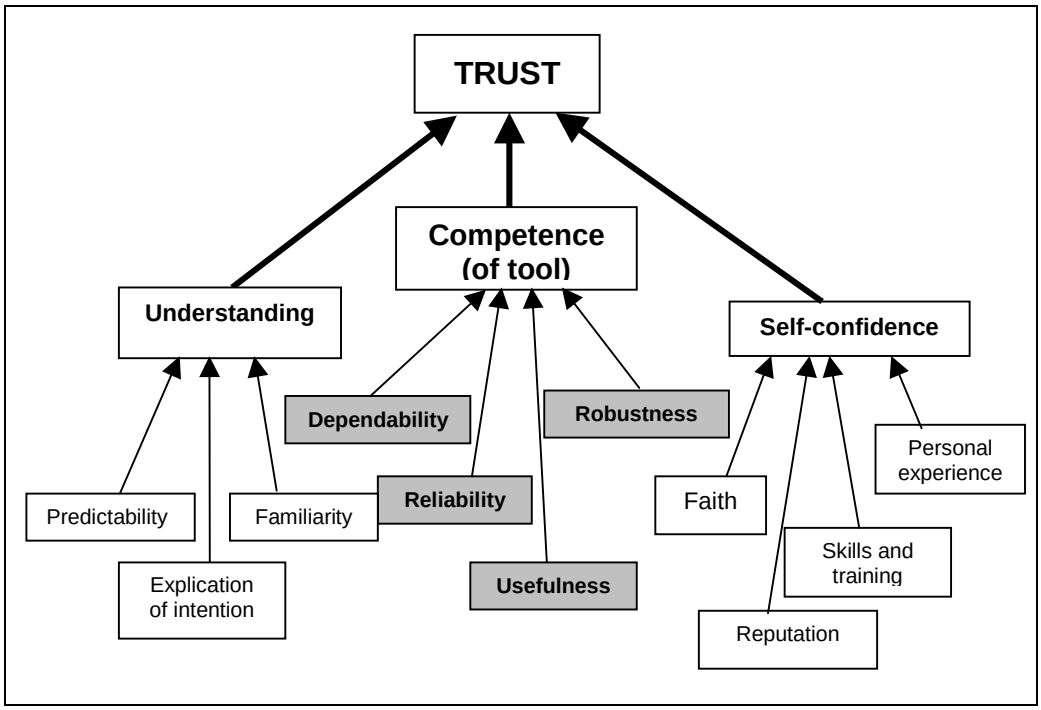

Figure 3: Simple model of trust and the relationship between factors

[4] See Web site http://www.tcd.ie/Psychology/aprg/proj.html

Page intentionally left blank

5. MEASURING TRUST

5.1 General

Given the requirement that the measure of trust must be usable in the context of real-time simulations, and hence be relatively easy to apply without being intrusive, the most appropriate measure *could* be a simple rating scale. A number of such scales are described in 5.2.

A more sophisticated measure of trust to be applied after, rather than during, a simulation run could be based on the well-known NASA-TLX workload measure. For example, ATCOs would rate their degree of trust on a number of dimensions (which would need to be determined) which are then summed to provide an overall score. Trust is not, of course, independent of the other proposed measures. That is, if a controller has good SA, if he/she is working well within the team, it is logical to assume that the controller has a high level of trust in the system being operated. Therefore, one could envisage that it is possible to combine, or embed, the measure of trust within another measure.

5.2 Subjective Measures

As stated earlier (see 3.1), the use of subjective rating scales is the most common means of measuring trust and four rating scales are described below[5].

1. Lee and Moray scale

   Lee and Moray (1992, 1994) employed a simple ten-point scale to evaluate operators' trust. The scale was administered after the end of each trial. In response to questions such as "Overall, how much do you trust the system?" the operators gave a score varying from 1 ('not at all') to 10 ('completely'). There are clear analogies to the use of ISA and the NASA TLX workload measures for ATC.

2. Muir scales

   Muir (1994), and Muir and Moray (1996) used a set of rating scales with the poles labelled 'none at all' or 'not at all' on the left, and 'extremely high' on the right. The operators were asked to rate their degree of trust in three aspects of a process control pump:

[5] It should be noted that subjective opinions do not always correspond to other measures that are simultaneously taken (e.g. Yeh & Wickens, 1984). If an operator says that he or she does or does not trust the system (or in the case of workload, feels overloaded by it), then other objective or physiological measures, whatever they imply, are difficult to interpret.

- your degree of trust in the pump to *respond* accurately,
- your degree of trust in the pump's *display,*
- your *overall degree of trust* in the pump.

In addition, the operators in this experiment also rated the pump's performance according to six other dimensions: competence, predictability, dependability, responsibility, reliability over time, and faith in future ability.

3. Madsen and Gregor

Drawing on the earlier work of Rempel *et al.* (1985), Sheridan (1988), Muir and Moray (1996), and others, Madsen and Gregor (2000) have developed a subjective measure for measuring trust of computers. The measure, called the Human-Computer Trust (HCT) scale, consists of five main constructs each with five sub-items as shown in Table 4. These five items are drawn from an original list of ten trust constructs as having the most predictive validity. Madsen and Gregor claim that the HCT has been empirically shown to be valid and reliable. The relationship between the five constructs is also shown in Figure 4.

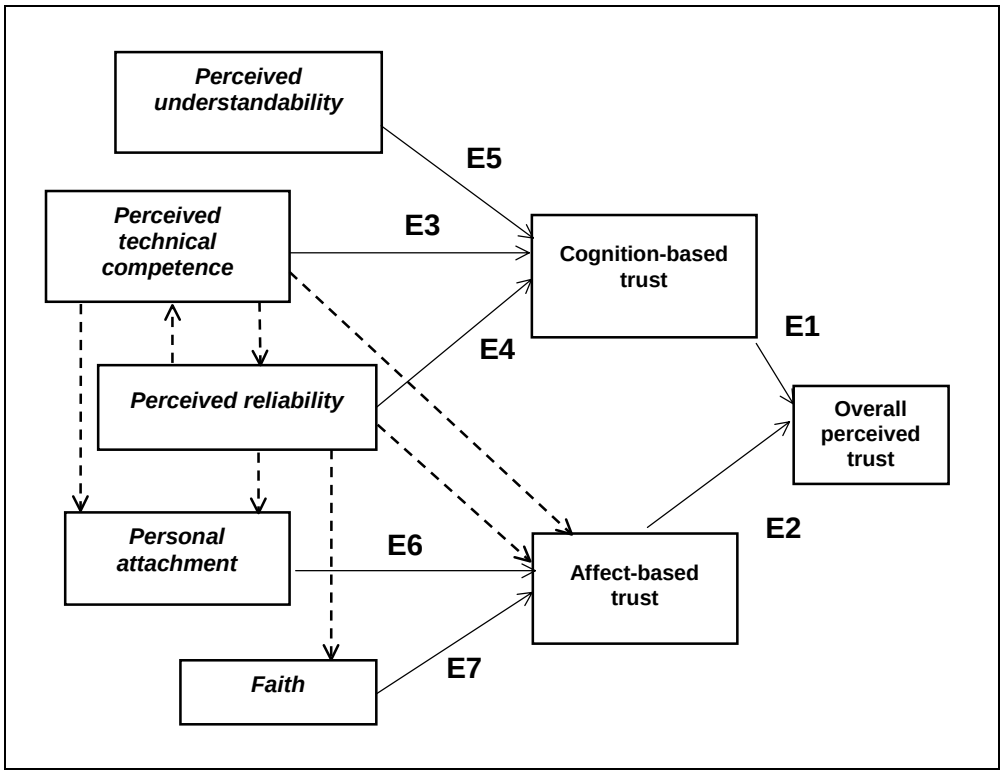

Figure 4: Model of Human-Computer Trust (HCT) components (from Madsen & Gregor, 2000)

Table 4: Human-Computer Trust (HCT) rating scale (Madsen & Gregor, 2000)

1. *Perceived reliability*
R1. The system always provides the advice I require to make my decision.
R2. The system performs reliably.
R3. The system responds the same way under the same conditions at different times.
R4. can rely on the system to function properly.
R5. The system analyzes problems consistently.
2. *Perceived technical competence*
T1. The system uses appropriate methods to reach decisions.
T2. The system has sound knowledge about this type of problem built into it.
T3. The advice the system produces is as good as that which a highly competent person could produce.
T4. The system correctly uses the information I enter.
T5. The system makes use of all the knowledge and information available to it to produce its solution to the problem.
3. *Perceived understandability*
U1. I know what will happen the next time I use the system because I understand how it behaves.
U2. I understand how the system will assist me with decisions I have to make.
U3. Although I may not know exactly how the system works, I know how to use it to make decisions about the problem.
U4. It is easy to follow what the system does.
U5. I recognize what I should do to get the advice I need from the system the next time I use it.
4. *Faith*
F1. I believe advice from the system even when I don't know for certain that it is correct.
F2. When I am uncertain about a decision I believe the system rather than myself.
F3. If I am not sure about a decision, I have faith that the system will provide the best solution.
F4. When the system gives unusual advice I am confident that the advice is correct.
F5. Even if I have no reason to expect the system will be able to solve a difficult problem, I still feel certain that it will.
5. *Personal attachment*
P1. I would feel a sense of loss if the system was unavailable and I could no longer use it.
P2. I feel a sense of attachment to using the system.
P3. I find the system suitable to my style of decision-making.
P4. I like using the system for decision-making.
P5. I have a personal preference for making decisions with the system.

4. Jian *et al*.

   Jian, Bizantz and Drury (2000) have developed a twelve-item trust 'questionnaire' incorporating a seven-point rating scale, where 1 on the scale equals 'not at all' and 7 equals 'extremely'. The trust questionnaire was developed as part of a three-phased experimental study. In the first phase, a word elicitation study, various words related to concepts of trust and distrust were collected (see Table 2). The second phase, a questionnaire study, investigated how closely each of these words was related to trust or distrust. The third phase was a paired comparison study, in which participants rated the similarity of pairs of words. Data from both the questionnaire study and the paired comparison study were then used to construct a multi-dimensional measurement scale for trust. The resultant scale is shown in Figure 5.

   Interestingly, in recent experimental studies conducted in the context of military command and control decision-making (Bisantz *et al.*, 2000), a computerised version of the trust questionnaire was employed.

5. Taylor *et al.*

   As part of extensive studies on the 'human-electronic crew' in the military domain, Taylor, Shadrake and Haugh (1995) a seventeen-item, seven-point rating scale questionnaire to determine operators views on the timeliness and appropriateness of adaptive computer aiding. The task environment consisted of three tasks, namely tracking, monitoring and resource management, which were carried out in various automation scenarios ('cooperative' and 'uncooperative'). The questionnaire is shown in Table 5.

6. Controller Acceptance Rating Scale (CARS)

   Another example is the Controller Acceptance Rating Scale (CARS) developed by researchers at the FAA (Lee & Davis, 1995). The scale was developed from the earlier Cooper-Harper scale (Cooper & Harper, 1969), but could from the basis of a measure of trust.

Below is a list of statement for evaluating trust between people and automation. There are several scales for you to rate intensity of your feeling of trust, or your impression of the system while operating a machine.
Please mark an 'x' on each line at the point which best describes your feeling or your impression.

(Note: 'not at all' = 1; 'extremely' = 7)

1. **The system is deceptive**

1 2 3 4 5 6 7

2. **The system behaves in an underhanded manner**

1 2 3 4 5 6 7

3. **I am suspicious of the system's intent, action, or outputs**

1 2 3 4 5 6 7

4. **I am wary of the system**

1 2 3 4 5 6 7

5. **The system's actions will have a harmful or injurious outcome**

1 2 3 4 5 6 7

6. **I am confident in the system**

1 2 3 4 5 6 7

7. **The system provides security**

1 2 3 4 5 6 7

8. **The system has integrity**

1 2 3 4 5 6 7

9. **The system is dependable**

1 2 3 4 5 6 7

10. **The system is reliable**

1 2 3 4 5 6 7

11. **I can trust the system**

1 2 3 4 5 6 7

12. **I am familiar with the system**

1 2 3 4 5 6 7

Figure 5: Checklist for trust between people and automation (Jian *et al.*, 2000)

Table 5: Trust and awareness scale (from Taylor *et al.*, 1995)

Construct	Description
1. Confidence	Confidence in own ability to successfully complete the tasks with the aid of the adaptive automation
2. Self-confidence	Confidence in own ability to successfully complete the tasks
3. Accuracy	Accuracy of own performance on the tasks with the aid of the adaptive automation
4. Self-accuracy	Accuracy of own performance on tasks
5. Automation confidence	Confidence in ability of the machine to support successful completion of the tasks
6. Automation accuracy	Accuracy of machine in supporting successful completion of tasks
7. Automation dependability	The extent to which you can count on the machine to provide the appropriate support to the tasks
8. Automation reliability	The extent to which you can rely on the machine to consistently support the tasks
9. Predictability	The extent to which you can anticipate and expect the machine to support the tasks
10. Risk	The probability of negative consequences of relying on the machine to support successful completion of the tasks
11. Impact / Survivability	The severity and criticality of adverse or negative consequences of relying on the machine to support successful completion of the tasks
12. Decision complexity	The extent to which the machines' decision on when and how to intervene and support the task can be regarded as a simple and obvious choice
13. Uncertainty / doubt	The extent to which you have confidence in the machines' decision on when and how to intervene and support the task
14. Judgement / awareness	The extent to which the machines' decision on when and how to intervene and support the task requires assessment, knowledge, and understanding of the task
15. Faith	The extent to which you believe that the machine will be able to intervene and support the tasks in other system states in the future
16. Demand for trust	Level of trust required from you when the machine intervenes and supports the task
17. Supply of trust	Level of trust actually provided by you when the machine intervenes and supports task

5.3 Objective Measures

The work of Moray and his colleagues (e.g. Moray, Inagaki & Itoh, 2000; Moray, 1999; Muir & Moray, 1996; Moray, Lee & Muir, 1995; Lee & Moray, 1992, 1994) has shown that this is possible to develop an empirical model of trust, and that the model equations are highly predictive. Studies such as these, at least within the context of supervisory control processes, suggest that trust in automation can:

a) be measured directly by asking the operators/controllers;

b) modelled on the basis of measurements of physical objective properties of the system in real-time;

c) modelled dynamically to predict trust, self-confidence and the probability of intervention by operators in automated systems.

Whether or not it is possible to apply such an approach to ATC has yet to be shown, but theoretically is entirely feasible. The particular set of variables relevant to ATC would have to be established empirically and subjected to sensitivity analyses. If trust is predictable from observable system physical properties (such as productivity output, selection of particular system functions, frequency of manual intervention, etc.) then the intervening variable (trust), even though not directly observable, becomes objectively measurable (see also 3.1).

A very simple, but crude measure could be whether or not the controller has activated a particular tool, the assumption being that if the tool has not been activated then it is not trusted. However, this is not necessarily true because it is possible that the system *is* trusted, but the controller thinks that he can do better than the automation. According to Moray *et al* (2000), it is known that there are cases where the human, or more often a combination of human and automation, exceed the performance as specified by a mathematically optimal automated system. In such circumstances manual control will be selected even thought the value of trust may be high.

It might be assumed that because a tool is open or activated the controller does indeed trust it. However, this is not necessarily true either, because the controller might simply be ignoring the information that is displayed by the tool. A more sophisticated measure is therefore needed. Consider the case of a conflict advisory tool such as MTCD. If it is used as intended one might expect that as soon as a conflict is displayed the controller issues instructions or enters flight data that is relevant to the conflict that has been detected. Both the type of data entered and the controller's speed of response could be measured and, theoretically at least, used to indicate the controller's level of trust.

5.4 Developing a Trust Measure for ATM Systems

Trust can be thought of as an 'enabler' to the introduction of new systems. It is useful therefore to measure controllers' trust during real time simulations. The development of a trust measure for evaluating automation support in ATM systems is properly the subject of a separate document (EATMP, 2003a). However, based on the literature that has been reviewed here, it is possible to provide some early recommendations.

The development of a subjective measure, using a rating scale, appears to be a simple and straightforward approach that has been used successfully in other domains. A rating scale to measure controllers' overall level of trust would seem to be an appropriate approach. Of the scales reviewed earlier (see 5.2) the one developed by Madsen and Gregor (2000) looks most promising particularly as the chosen constructs have shown to have a degree of empirical validity. The scale developed by Jian, Bisantz and Drury (2000) on the other hand would seem to be rather emotive. The use of terms like 'deceptive', 'suspicious' and 'underhand manner' (see Figure 5) runs the risk of implanting ideas of untrustworthiness where none previously existed.

As mentioned earlier (see 3.4 and 4.2) there is some evidence that controllers do not necessarily think in terms of degrees of 'trust' *per se*, but rather are concerned with the operational reliability of the tools that they use. They either both trust and use an ATM system, or they do not trust it all. Therefore, as an alternative to a single rating scale to measure overall trust, a set of rating scales to measure different *dimensions* of trust or confidence might prove more beneficial. Of the dimensions identified (see 3.2) the most promising (i.e. that are appropriate to ATM automation and minimise the possible ambiguity of terms) are:

- reliability,
- accuracy,
- understanding,
- faith,
- liking,
- familiarity,
- robustness.

The question of whether or not trust is inherently all-or-none is an interesting one. It may be that fuzzy set measures would be more appropriate, since the fuzzy set operators often behave like a switch despite the underlying variables being continuous. (For example, "Do you think the system is completely reliable, quite reliable, not reliable …?"; "Do you trust the system completely, quite a lot, not very much …?")

6. TRUST SUBJECTS

6.1 Introduction

It is evident from the literature reviewed that most of the research data on human trust of automation has been gathered from simulations of process control systems. Process control has some similarities with ATC in that they are both safety critical systems. In particular, the research has focused upon how trust is affected by the presence of system faults. Little research has been conducted to address controllers' trust of ATM systems, either in general or specifically of new automation tools.

Research into ATM systems has tended to be focused upon measuring controllers' *workload* when using certain tools, and ascertaining controllers' attitudes to those tools. The subject of trust has arisen indirectly as part of the explanations as to why controllers did, or did not, use the tools as expected.

That being said, on the basis of the research reviewed and analysed in this document a number of useful subjects to be considered for facilitating and promoting trust in the design and development of ATC systems can be given. These guidelines are presented below.

6.2 Subject - Terminology

1. Automation is a device or system that accomplishes (partially or fully) a function that was previously carried out (partially or fully) by a human operator.

2. Trust is the extent to which a user is confident in and willing to act on the basis of, the recommendations, actions, and decisions of a computer-based 'tool' or decision aid.

3. Trust is a construct composed of different elements or dimensions. The main dimensions identified in the research literature are, for example:

   - predictability,
   - reliability,
   - dependability,
   - understandability,
   - self-confidence.

4. Evidence from many empirical studies, in areas as diverse as inter-personal psychology and supervisory process control, indicates that trust is not a discrete, binary (yes/no) variable, but varies on a continuum from no trust to complete trust.

5. Contrary to the last point, anecdotal evidence from simulations and other ATC environments suggests that although controllers have varying levels of *confidence* in their equipment, they do in fact perceive trust as a discrete variable. If they trust something, they will use it provided that their level of confidence is above a certain criterion level (defined by experience); if they do not trust something, they will not use it. This observation needs to be borne in mind when discussing the issue of trust with controllers.

6.3 Subject - Trust in Automation

6. It is NOT the case that the object of training and experience is to "make the controllers trust the system". It is for them to develop trust *at an appropriate level*, neither too much nor too little. There are cases where training should make the controller not trust the system, i.e. when it is not reliable. (This is similar to the notion that a system should "match the mental model of the operator". Yes, but only if the mental model of the operator is correct!)

7. To foster appropriate trust in automation tools it is essential that the intended users (controllers) properly understand the purpose and functionality of the tools they are to use.

8. In order to understand the automation tool(s), controllers must be adequately trained. This might appear an obvious statement to make, but anecdotal evidence from many ATM system simulations suggests that controllers are not always taught the full functionality of the tools that they are expected to use.

9. It should *not* be assumed that because controllers have completed a preliminary few days of training that they understand how a particular tool functions; this should be tested.

10. It is very important that controllers understand why, and under what conditions, an automation tool might make errors. Leaving aside the problems of testing a prototype tool in a simulation, even if the tool is working as intended it is unrealistic to expect it to be perfect. Controllers need to be aware of the problems. Trust will grow if operators find compensating strategies for the consequences of an automation error.

    STCA provides a good example of the likely problems to be encountered. Even though STCA is now in operational service, it is known to have imperfections - nuisance alerts and other spurious false alarms. The development of an automation tool such as MTCD will inevitably suffer the same problems. It is unrealistic to expect otherwise and controllers should be briefed about the known limitations of the tool.

11. As a corollary of the above, in order to promote the building of trust in a system the opportunity for the system to 'miss' should be made as low as possible. However it is often the case that by increasing the sensitivity of a system to improve detection rates the number of false alarms are increased. This may not appear to be a problem, however false alarms give the impression of incompetence and are therefore likely to erode trust in the system, demonstrated by operators ignoring or disabling these systems. It is vital therefore that the constraints of the system are carefully considered to instil maximum trust.

12. Trust is strongly affected by system reliability (as one would expect) but self-confidence is not (at least in a system in which operators can distinguish the tasks they perform manually from those performed by automation and in which it is the latter that are mainly affected by unreliability). It is important therefore that the system is reliable before commencing any assessments of performance.

13. There is sometimes a tendency in those developing new automated (or other) systems that are safety-critical or involve high hazard, to under-estimate failure likelihood of those systems. Experience shows that all systems can fail, and often failure rates are much higher during the initial introductory periods with new systems, despite extensive pre-operational testing. We KNOW that they can fail, so it is therefore essential to prepare both the system (hardware and software) and the controller for the eventuality of some form of system failure.

14. The causal factors driving the dynamics of trust are different from those driving self-confidence. Trust seems to be reduced by *properties* of the system (real or apparent false diagnoses), whereas self-confidence is reduced by experiences of the operator (experiences of accidents). High self-confidence often produces a bias in favour of manual control.

15. Experimental evidence indicates that *distrust* is more resistant to change than trust. In practical terms this suggests that if controllers lose trust in the automated tools that they are using, it is difficult to regain that trust. This reinforces the message that simulations with new automated tools should not be undertaken unless a reasonable degree of system reliability can be guaranteed (and that the controllers have been briefed about known limitations of the tools).

16. There is some empirical evidence to indicate that operators' trust of 'intelligent' automation (i.e. knowledge-based systems, decision support systems, expert systems, etc.) is not the same as their trust of simpler, automated functions (e.g. conflict alerts). This has implications for the development of ATM automation that is intended to advise the controller on courses of action. That is, the need to ensure that controllers understand the automation, that they are properly trained, and that the system is reliable, will be even more important.

6.4 Subject - Measuring Trust

17. Trust can be thought of as an 'enabler' to the introduction of new systems. It is useful therefore to measure controllers' trust during real time simulations. Evidence from many empirical studies indicates that a subjective, questionnaire-based technique will be most appropriate.

18. Measuring controllers' trust in the context of real time simulations (with all of the inherent problems that simulations bring) requires a technique that is simple and straightforward to apply. For the purposes of the SHAPE project it is recommended that a set of rating scales be used that measure both the controllers' overall level of trust, and the constituent elements of trust or confidence (e.g. reliability, predictability, understandability, etc.). The latter type of measure is important because anecdotal evidence suggests that controllers are not entirely 'comfortable' with the concept of trust. That is, they tend to view the equipment and systems that they operate in terms of its operational reliability.

19. When constructing the trust measure care needs to be taken about the exact words that are employed in the questionnaire, rating scales or other measure. As shown in several empirical studies and from anecdotal evidence from discussions with controllers, words such as 'reliability', 'accuracy', and even 'trust', mean different things to different people. This is especially important to consider when the intended recipients of the trust measure are not native speakers of English (as is the case with ATCOs).

20. Evidence from empirical studies indicates that the concepts of trust and distrust can be treated as opposite ends of a trust continuum. In practical terms this means that a single scale (e.g. from -5 to +5) could be used to measure levels of trust.

21. The question of whether or not trust is inherently all-or-none is an interesting one. It may be that fuzzy set measures would be more appropriate, since the fuzzy set operators often behave like a switch despite the underlying variables being continuous.

22. In order to collect a useful amount of data about trust, the measure will need to be given to controllers on repeated occasions, probably after each simulation run, or at least once every day.

Controllers' trust in automation is a key determinant in the development and implementation of new ATM systems. In order to develop that trust at an appropriate level, and avoid inappropriate distrust, it is essential that:

- **controllers understand the functionality of the automation, and its limitations;**
- **controllers are given proper and sufficient training;**
- **the simulation system in general and the automation in particular are highly reliable.**

Page intentionally left blank

REFERENCES

Abdul-Rahman, A. & Hailes, S. (1999). Relying on trust to find reliable information. *1999 International Symposium on Database, Web and Cooperative Systems (DWACOS'99)*, Baden-Baden, Germany.

Abdul-Rahman, A. & Hailes, S. (2000). Supporting trust in virtual communities. *Hawaii International Conference on System Sciences 33*, Maui, Hawaii, 4-7 January 2000.

Bainbridge, L. (1982). Ironies of automation. In: G. Johannsen & J.E. Rijnsdorp*, Analysis, Design and Evaluation of Man-Machine Systems*, Proc. of IFAC Conf., Baden-Baden, Germany, 129-135.

Bisantz, A.M., Llinas, J., Seong, Y., Finger, R. & Jian, J.Y. (2000). *Empirical investigations of trust-related system vulnerabilities in aided, adversarial decision-making*. Center for Multi-Source Information Fusion, Dept. of Industrial Engineering, State University of New York at Buffalo.

Billings, C.E. (1991). Human-centred Aircraft Automation Philosophy: A Concept and Guidelines. *NASA Technical Memorandum No. 103885*. National Aeronautics and Space Administration.

Bonini, D., Jackson, A. & McDonald, N. (2001). Do I trust thee? An approach to understanding trust in the domain of air traffic control. In: *Proceedings of People in Control*, 19-21 June, UMIST Manchester.

Cardosi, K. & Murphy, E. (1995). *Human Factors in the Design and Evaluation of Air Traffic Control Systems*. Federal Aviation Administration, Office of Aviation Research, DOT/FAA/RD-95/3.

Chabrol, C., Vigier, J.C., Garron, J. & Pavet, D. (1999). CENA PD/3 Final Report, PHARE/CENA/PD/3-2.4/FR/2.0.

Cooper, G.E. & Harper, R.P. (1969). *The use of pilot rating in the evaluation of aircraft handling qualities*, NASA-AMES Report TN-D-5153.

Cox, M. & Kirwan, B. (1999). The future role of the air traffic controller: Design principles for human-centred automation. In: M.A. Hanson, E.J. Lovesey & S.A. Robertson, *Contemporary Ergonomics 1999*, Taylor & Francis Ltd., 27-31.

DERA (1997) WP6: Application of evaluation techniques. Annex B. Results of DERA cognitive walkthrough activity. EC DGVII RHEA Project, Ref. RHEA/TH/WPR/6/2.0, 30th July.

Du Boulay, E., Cox, M., Hawkins, J.R. & Williams, J. (1994). NODE-M STCA: ATC Evaluation. National Air Traffic Services, CS Report 9439, May.

Dzindolet, M., Pierce, L.G., Beck, H.P. & Dawe, L. (1999). Misuse and disuse of automated aids. Proc. of the Human Factors and Ergonomics Society 43rd Annual Meeting, 339-343.

Dzindolet *et al.* (2000a). Building trust in automation. Paper presented at *Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millenium*, The 4th Conference on Automation Technology and Human Performance and the 3rd Conference on Situation Awareness in Complex Systems, October 15-19.

Dzindolet, M., Pierce, L.G., Beck, H.P. & Dawe, L. (2000b). A framework of automation use. *Manuscript submitted for publication*.

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

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

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

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

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

Erzberger, H. (1989). ATC automation concepts. In: *Proceedings of the Aviation Safety and Automation Program Conference*, NASA Conf. Publication 3090.

EUROCONTROL (1993). *Role of Man within PHARE*; EUROCONTROL DOC 93-70-35.

EUROCONTROL (2000a). Air traffic controller attitudes toward future automation concepts: A literature review. In: EUROCONTROL Report ASA.01.CORA.2.DEL02-A.RS, 4th December.

EUROCONTROL (2000b). Conflict Resolution Assistant level 2 (CORA2). Controller assessments. In: EUROCONTROL Report ASA.01.CORA.2. DEL02-b.RS, 4th December.

Funk, K., Lyall, B. & Riley, V. (1996). A comparative analysis of flightdecks with varying levels of automation. In: *Phase 1 Final Report: Perceived human factors problems of flight deck automation*, FAA.

Goillau, P., Woodward, V., Kelly, C. & Banks, G. (1998). Evaluation of virtual prototypes for ATC – the MACAW technique. In: M. Hanson (Ed) (1998) *Contemporary Ergonomics '98*, London: Taylor & Francis, p. 419-423.

Graham, R., Young, D., Pichancourt, I., Marsden, A. & Irkiz, I. (1994). ODID IV simulation report. *EEC Report No. 269/94*. Brétigny-sur-Orge, France: EUROCONTROL.

Hale, S. & Baker, S. (1990). The presentation of Short Term Conflict Alert: A human factors perspective. Civil Aviation Authority, *DORA Report 9018*, June.

Hall, R.J. (1996). Trusting your assistant. In: *Proceedings of KBSE '96*, 42-51.

Hollnagel, E., Cacciabue, P.C. & Bagnara, S. (1994). Workshop report. The limits of automation in air traffic control; *Int. J. Human-Computer Studies*, 40, 561-566.

Hopkin, V.D. (1975). The controller versus automation. In: AGARD AG-209.

Hopkin, V.D. (1995). *Human Factors in Air Traffic Control*. Taylor & Francis Ltd.

Hopkin, V.D. (1998). The impact of automation on air traffic control specialists. In: M.W. Smolensky & E.S. Stein, *Human Factors in Air Traffic Control*, Academic Press, 391-419.

ICAO (1994). *Human Factors Digest No.11. Human factors in CNS/ATM systems. The development of human-centred automation and advances technology in future aviation systems*. International Civil Aviation Organization, ICAO Circular 249-AN/149.

Inagaki, T. (1999). Automation may be given the final authority. Proceedings of CybErg 1999: The 2nd Int. Cyberspace Conf. on Ergonomics. Int. Ergonomics Assoc. Press, 68-74.

Jian, J.-J., Bisantz, A.M. & Drudy, C.G. (1998). Towards an empirically determined scale of trust in computerized systems: Distinguishing concepts and types of trust. Proc. of the Human Factors and Ergonomics Society Annual Meeting, Chicago, 501-505.

Jian, J.-J., Bisantz, A.M. & Drudy, C.G. (2000). Foundations for an empirically determined scale of trust in automated systems. *Int. J. of Cognitive Ergonomics*, 4(1), 53-71.

Kelly, C.J., Goillau, P.J., Finch, W. & Varellas, M. (1995*). CAER Future System 1 (FS1) Final trial report*. Defence Research and Evaluation Agency, Report No. DRA/LS(LSC4)/CTR/RPT/CD246/1.0, November.

Kelly, C.J. & Goillau, P.J. (1996). Cognitive Aspects of ATC: Experience of the CAER & PHARE Simulations; Paper presented at *8th European Conference on Cognitive Ergonomics (ECCE - 8),* Granada, 10th - 13th September.

Kramer, R.M. (1999). Trust in organizations: Emerging perspectives, enduring questions. Annu. Rev. Psychology, Vol 50, 569-598.

Lee, K. & Davis, T.J. (1995). The development of the Final Approach Spacing Tool (FAST): A cooperative controller-engineer design approach; NASA Technical Memorandum 110359, August.

Lee, J.D. & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, **35**, 10, 1243-1270.

Lee, J.D. & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. *Int. J. Human-Computer Studies*, **40**, 153-184.

Lewandowsky, S., Mundy, M. & Tan, G. (2000). The dynamics of trust: Comparing humans to automation. *J. of Experimental Psychology: Applied*. Vol 6, 2, 104-123.

Liu, C. & Hwang, S. (2000). Evaluating the effects of situation awareness and trust with robust design in automation. *International Journal of Cognitive Ergonomics*, 4 (2), 125-144.

Madsen, M. & Gregor, S. (2000). Measuring human-computer trust. In: *Proceedings of Eleventh Australasian Conference on Information Systems*, Brisbane, 6-8 December.

Masalonis, A.J., Duley, J., Galster, S., Castano, D., Metzger, U. & Parasuraman, R. (1998). Air traffic controller trust in a conflict probe during Free Flight. Proc. of the 42nd Annual meeting of the Human Factors and Ergonomics Society, 1607.

Moffa, A.J. & Stokes, A.F. (1997). Trust in a medical system: Can we generalize between domains? In: M. Mouloua & J.M. Koonce (Eds). *Human-Automation Interaction: Research and Practice*. Mahwah, New Jersey: Lawrence Erlbaum Associates, pp 127-224.

Moray, N. (1999). Monitoring, complacency, scepticism and eutactic behaviour. In: *Proceedings of CybErg 1999: The 2nd Int. Cyberspace Conf. on Ergonomics*. Int. Ergonomics Assoc. Press.

Moray, N., Inagaki, T. & Itoh, M. (2000). Adaptive automation, trust and self-confidence in fault management of time-critical tasks; *J. of Experimental Psychol: Applied*, **6**, 1, 44-58.

Moray, N., Inagaki, T. & Parasuraman, R. (2001). Attention and complacency. *Paper submitted for publication*.

Moray, N., Lee, J. & Muir, B. (1995). Trust and human intervention in automated systems. In: J.M. Hoc, P. Cacciabue & E. Hollnagel (Eds*). Expertise in Technology. Cognition and Human-Computer Cooperation*. Lawrence Erlbaum Associates.

Mouloua, M. & Koonce, J. (1997). *Human-automation interaction: Research and Practice*. Lawrence Erlbaum Associates.

Muir, B. (1987). Trust between humans and machines, and the design of decision aids, *Int. J. Man-Machine Studies*, 27, 527-539.

Muir, B. (1994). Trust in automation: Part 1. Theoretical issues in the study and human intervention in automated systems. *Ergonomics*, **37**, 1905-1923.

Muir, B. & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, **39**, 3, 429-460.

Nijhuis, H. (1998). Automation Philosophies: Evaluation of Some Concepts. In: *Proceedings of the Third EUROCONTROL Human Factors Workshop: Integrating Human Factors into the Life Cycle of ATM Systems*. Luxembourg, 7-9 October 1998. HUM.ET1.ST13.000-REP-03. Edition 1.0 Released Issue. Brussels: EUROCONTROL.

Nijhuis, H., Buck, S., Kelly, C., Goillau, P., Fassert, C., Maltier, L. & Cowell, P. (1999). WP8: Summary and consolidation of RHEA results. European Commission DGVII, Report RHEA/NL/WPR/8/04, 28th Feb.

NRC (1997). *More Than Screen Deep: Toward Every-Citizen Interfaces to the Nation's Information Infrastructure*. Commission on Physical Sciences, Mathematics and Applications. National Research Council.

Parasuraman, R. & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, **39**, 2, 230-253.

Parasurman, R., Molloy, R. & Singh, I.L. (1993). Performance consequences of automation-induced "complacency". *Int. J. of Aviation Psychology*, 3, 1-23.

Parasuraman, R., Sheridan, T.B. & Wickens, C. (2000). A model for types and levels of human interaction with automation. *IEEE Trans. On Systems, Man and Cybernetics-Part A: Systems and Humans*, Vol 30, No.3, 286-297.

Reichmuth J., Schick, F., Adam, V., Hobein, A., Link, A., Teegen, U. & Tenoort*, S.* (1998). PD/2 Final Report. EUROCONTROL PHARE Report PHARE/DLR/PD/2-10.2/SSR;1.2. February.

Rempel, J.K., Holmes, J.G. & Zanna, M.P. (1985). Trust in close relationships. J. of Personality and Social Psychology, 49, 1, 95-112.

Riley, V. (1994). Human use of automation. Unpublished doctoral dissertation, University of Minnesota.

Scerbo, M.W. & Mouloua, M. (1999). Automation technology and automation performance: Current research and trends. Lawrence Erlbaum Associates.

Schneider, F.B. (1999). Trust in cyberspace. National Academy Press.

Sheridan, T.B. (1988). Trustworthiness of command and control systems. *Proc. of Analysis, Design and Evaluation of man-Machine Systems 1988*, 3rd IFAC/IFIP/IEA/IFORS Conf., Finland, 14-16 June.

Shorrock, S. & Scaife, R. (2000). Evaluation of an alarm management system for an ATC centre. In: D. Harris (Ed). *Engineering Psychology and Cognitive Ergonomics: Volumes 5 and 6*. Ashgate Publishing.

Simpson, A. (1992). *HCI issues in trust and acceptability*; Defence Evaluation and Research Agency, Report No. DRA TM(CAD5) 92018, November.

Simpson, A. (1995). Seaworthy trust: Confidence in automated data fusion. In: R. Taylor & J. Reising (Eds). *The Human-Electronic Crew: Can we Trust the Team? Proc. of the 3rd Int. Workshop on Human-Computer Teamwork.* Defence Evaluation and Research Agency, Report No. CHS/HS3/TR95001/02, 77-81.

Stoner, C. (1995). Controllers as air traffic managers. In: *Proceedings of Global NAVCOM'95*, Montreal, 23-25 May.

Tan, G. & Lewandowsky, S. (1996). A comparison of operator trust in humans versus machines. Proc. of CybErg 1996: The 1st Int. Cyberspace Conf. on Ergonomics. Int. Ergonomics Assoc. Press.

Taylor, R.M. (1988). Trust and awareness in human-electronic crew teamwork. In: *The Human-Electronic Crew: Can They Work Together?* Wright-Patterson AFB, OH., Report WRDC-TR-89-7008.

Taylor, R.M., Shadrake, R. & Haugh, J. (1995). Trust and adaptation failure: An experimental study of uncooperation awareness. R. Taylor & J. Reising (Eds), *The Human-Electronic Crew: Can we Trust the Team? Proc. of the 3rd Int. Workshop on Human-Computer Teamwork.* Defence Evaluation and Research Agency, Report No. CHS/HS3/TR95001/02, 93-98.

Weigner, M.B. (1997). Human-user medical device interactions in the anesthesia work environment. In: M. Mouloua & J. Koonce (Eds), *Human-Automation Interaction: Research and Practice*. Lawrence Erlbaum Associates, 241-248.

Whitaker, R. & Marsh, D. (1997) PD/1 Final Report, PHARE Report DOC 96-70-24, PHARE/NATS/PD1-10.2/SSR, 1.1.

Wickens, C.D., Mavor, A.S. & McGee, P. (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.

Wiener, E.L. (1985). Beyond the sterile cockpit. *Human Factors*, **27**, 1, 75-90.

Wiener, E.L. & Curry, R.E. (1980). Flightdeck automation: promises and problems. Ergonomics, 23, (10), 995-1011.

Wise, J.A., Hopkin, V.D. & Smith, M.L. (1991). *Automation and system issues in air traffic control*. Springer-Verlag.

Yeh, Y.Y. & Wickens, C.D. (1984). Why do performance and subjective workload measures dissociate? In: *Proceedings of the Human Factors Society 28th Annual Meeting*, 504-508.

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
CAER	Computer Assistance for En-Route ATC
CARS	Controller Acceptance Rating Scale
CBT	Computer-Based Training
CORA1/2/3	Conflict Resolution Assistant 1/2/3
DERA	Defence Evaluation and Research Agency *(UK; now known as QinetiQ)*
DIS	Director(ate) Infrastructure, ATC Systems and Support *(EUROCONTROL Headquarters, SDE)*
DIS/HUM	See 'HUM (Unit)'
EATCHIP	European Air Traffic Control Harmonisation and Integration Programme *(now EATMP)*
EATMP	European Air Traffic Management Programme *(formerly EATCHIP)*
EEC	EUROCONTROL Experimental Centre *(Brétigny, France)*
ECI	Every-Citizen-Interface
FAA	Federal Aviation Administration *(US)*
FRAP	Free Route Airspace Project
FREER	Freer Flight
GUI	Guidelines *(EATCHIP/EATMP)*
HCT	Human-Computer Trust

HFSG	Human Factors Sub-Group *(EATCHIP/EATMP, HUM, HRT)*
HIPS	Highly Interactive Problem Solver
HRS	Human Resources Programme *(EATMP, HUM)*
HRT	Human Resources Team *(EATCHIP/EATMP, HUM)*
HSP	Human Factors Sub-Programme *(EATMP, HUM, HRS)*
HUM	Human Resources (Domain) *(EATCHIP/EATMP)*
HUM (Unit)	Human Factors and Manpower Unit *(EUROCONTROL Headquarters, SDE, DIS; also known as 'DIS/HUM'; formerly stood for the 'ATM Human Resources Unit')*
IANS	Institute of Air Navigation Services *(EUROCONTROL, Luxembourg)*
ISA	Instantaneous Self-Assessment
MTCD	Medium-Term Conflict Detection
NASA	National Aeronautics and Space Administration *(US)*
NERC	New En-Route Centre
NLR	Nationaal Lucht- en Ruimtevaartlaboratorium *(National Aerospace Laboratory, NL)*
NRC	National Research Council
NTT	NERC Transition Team
ODID	Operational Display and Input Development
PD/1/2/3	PHARE Demonstration 1/2/3
PHARE	Programme for Harmonised Air Traffic Management Research in EUROCONTROL
REP	Report *(EATCHIP/EATMP)*
RHEA	Role of the Human in the Evolution of ATM systems
SDE	Senior Director, Principal EATMP Directorate *or, in short,* Senior Director(ate) EATMP (*EUROCONTROL Headquarters)*

SHAPE (Project)	Solutions for Human-Automation Partnerships in European ATM (Project) *(EATMP, HUM, HRS, HSP)*
STCA	Short-Term Conflict Alert
TLX	Task Load Index

Page intentionally left blank

ACKNOWLEDGEMENTS

The contribution of the Members of the HRT Human Factors Sub-Group to this document during the meetings, and further written comments, were much appreciated.

Neville Moray, and Barry Kirwan[6] from the EUROCONTROL Human Factors and Manpower Unit, provided comments on the deliverable, and their help is gratefully acknowledged.

Document Configuration

Carine Hellinckx
(External contractor)
EUROCONTROL Headquarters, DIS/HUM

[6] Now works at the EEC

Page intentionally left blank