



Executive summary

The influence of automation support on performance, workload and situation awareness of Air Traffic Controllers



Report no.

NLR-TP-2009-700

Author(s)

G.K. van de Merwe
E.A.P.B. Oprins
D.J.F. Eriksson
A.H. van der Plaats

Report classification

UNCLASSIFIED

Date

February 2010

Knowledge area(s)

Luchtverkeersmanagement(ATM)-
en luchthavensimulatie en -
validatie
Training, Simulatie en Operator
Performance

Descriptor(s)

SARA
Situation Awareness
Workload
Automation support
Air Traffic Control

Problem area

A time-based operation, as planned in the ATM future, is assumed to affect the controllers' Situation Awareness (SA) due to a higher priority of meeting a time objective and increasing automation. LVNL's future ATM system requires an improved punctuality at the Initial Approach Fix (IAF) to enable Continuous Descent Approaches (CDAs) in the Schiphol TMA. This

paper provides SA requirements on the design of controller support tools in time-based operations, based on a short literature review and an empirical study (Real-Time Simulation and an Operational Trial) executed at Air Traffic Control the Netherlands (LVNL).

Description of work

In order to address these issues the SARA concept (Speed And Route

This report is based on a paper to be published in the International Journal of Aviation Psychology.

Advisor) was tested in a Real-Time Simulation (RTS) and an Operational Trial. An RTS was chosen to evaluate the concept in a controlled experimental environment. Subsequently the Operational Trial was executed to further evaluate the concept in a real-life work environment. In both experiments, the influence of SARA on controllers' performance, workload and SA was investigated.

Results and conclusions

The SARA real-time experiment and the Operational Trial results showed that this tool does not increase the controllers' workload (R/T load, inputs), while the target of a higher accuracy at IAF was met. The findings have also pointed at two major impacts on the controllers' SA as expected from the literature. First, controllers are currently focusing more on distance than on time in forming a mental picture of the traffic situation. This changes their working strategies in sequencing traffic and solving conflicts. Second, additional automation (cf. SARA advisories)

could be in conflict with the controllers' own plan of traffic handling. They could lose a certain 'feeling of control' and ultimately their SA. However, there was a strong learning effect already after a few experimental sessions. This suggests that a gradual implementation and training will certainly help supporting a smooth introduction. Moreover, the impact on SA appears to depend on the specific design (e.g. Human Machine Interface (HMI), separation responsibility, quality of advisories).

Applicability

The specific design of the tool and the controllers' familiarity with it determine the degree to which a sufficient SA could be maintained. Therefore the responsibilities between the humans and the systems should be carefully assessed when designing and implementing automation solutions.



NLR-TP-2009-700

The influence of automation support on performance, workload and situation awareness of Air Traffic Controllers

G.K. van de Merwe, E.A.P.B. Oprins¹, D.J.F. Eriksson² and
A.H. van der Plaat²

¹ Luchtverkeersleiding Nederland (LVNL), currently employed at TNO Defence and Security

² Luchtverkeersleiding Nederland (LVNL)

This report is based on a paper to be published in the International Journal of Aviation Psychology.
The contents of this report may be cited on condition that full credit is given to NLR and the authors.
This publication has been refereed by the Advisory Committee AIR TRANSPORT.

Customer National Aerospace Laboratory NLR
Contract number ----
Owner National Aerospace Laboratory NLR
Division NLR Air Transport
Distribution Unlimited
Classification of title Unclassified
April 2010

Approved by:

Author	Reviewer	Managing department
i.o.  27-5-'10	 27-5-'10	 27-5-'10

Summary

A decision support tool, called Speed And Route Advisor (SARA), was developed at Schiphol Airport to help air traffic controllers with achieving an increased accuracy in traffic delivery. Its influence on controller performance, workload and Situation Awareness (SA) was evaluated in a Real-Time Simulation and in an Operational Trial. The findings indicate that this additional system support is necessary to achieve higher accuracy without increasing the controllers' workload. At the same time, controllers must stay in-the-loop to maintain SA. This must be kept in mind while designing decision support systems such as SARA.

Contents

1	Introduction	7
2	Work complexity in ATM	7
2.1	ATC performance	7
2.2	Workload management	9
2.3	Situation awareness	9
3	Automation in ATM	9
3.1	Workload	10
3.2	Situation Awareness	10
3.3	Trust	11
4	Future ATM developments	11
4.1	Arrival Management	12
4.2	Arrival management at Schiphol Airport	12
5	SARA 13	
5.1	Basic functioning	13
5.2	The impact of SARA on air traffic controllers	14
6	Study 1: Real Time Simulation	15
6.1	Experimental design	15
6.2	Results	17
6.2.1	EAT adherence	17
6.2.2	Workload	18
6.2.3	Situation Awareness	21
7	Study 2: Operational Trial	22
7.1	Experimental design	22
7.2	Results	23
7.2.1	EAT adherence	23
7.2.2	Workload	23
7.2.3	Situation Awareness	24



8	General Discussion	25
8.1	Performance	25
8.2	Workload	26
8.3	Situation Awareness	27
8.4	Future directions: system support in Time-Based Operations	28
8.5	Conclusion	29
9	Acknowledgements	29
10	References	30

Abbreviations

ACC	Area Control Centre
ANOVA	Analysis of Variance
ANSP	Air Navigation Service Provider
AoR	Area of Responsibility
APP	Approach
ATC	Air Traffic Control
ATM	Air Traffic Management
EAT	Estimated Arrival Time
EHAM	Schiphol Airport
ETO	Estimated Time Over
FMC	Flight Management Computer
HMI	Human Machine Interface
HFI	Human Factors Indication
IAF	Initial Approach Fix
ISA	Instantaneous Self Assessment
LVNL	Luchtverkeersleiding Nederland
MUAC	Maastricht Upper Area Control
NARSIM	NLR ATC Research SIMulator
NLR	Nationaal Lucht- en Ruimtevaartlaboratorium
R/T	Radio/Telephony
RTS	Real-Time Simulation
SARA	Speed and Route Advisor
SA	Situation Awareness
SASHA	Situation Awareness for SHAPE
SESAR	Single European Sky ATM Research
TID	Touch Input Device

1 Introduction

This article describes the influence of a decision support tool named Speed And Route Advisor (SARA) on the performance, workload, and Situation Awareness (SA) of air traffic controllers. Similar to other process control tasks in transportation (aviation, shipping, railways) or process industry (e.g., chemical and nuclear plants), the air traffic control (ATC) task is considered highly complex and dynamic (Oprins, 2008). Complex cognitive processes are required to handle the large amount of dynamically changing information in a three-dimensional environment (Garland, Stein and Muller, 1999). Therefore, ATC is also called a complex cognitive or high-performance skill (Schneider, 1990).

Air traffic is expected to grow at many airports, also at Schiphol Airport. At the same time, workload for air traffic controllers may not increase in the future because this might cause problems with keeping a sufficient number of controllers competent at their tasks. Much ATM research (e.g. SESAR, 2007; NextGen, 2007) focuses on the design of decision support tools that make it possible for controllers to handle larger amounts of traffic with reduced workload. At Air Traffic Control the Netherlands (cf. Luchtverkeersleiding Nederland; LVNL), a Speed And Route Advisor was developed (Oprins, Zwaaf, Eriksson, Merwe and Roe, 2009; Merwe, Oprins, Plaat, and Eriksson., 2009). This decision support tool provides controllers with a speed and a route advice, with which a higher punctuality of flights can be achieved whilst keeping the workload of controllers at an acceptable level. However, a potential risk is the possible decrement in the controllers' SA as shown in previous research on automation of ATM systems (Endsley, 1997; Metzger, 2001; Metzger and Parasuraman, 2005). The human factor impacts on performance, workload and SA must be addressed and evaluated when designing new decision support tools.

2 Work complexity in ATM

2.1 ATC performance

Due to the complex cognitive nature of the ATC task only a small number of people are able to acquire the required competences within a reasonable period of training (Schneider, 1990). LVNL is coping with a shortage of controllers which is not uncommon at busy and complex airports. LVNL is attempting to solve this problem by improving selection and training, and by designing new ATM concepts that make the work less complex. As a starting point, a competence analysis was performed at LVNL based on literature research and workshops with controllers. This has resulted in the so-called ATC Performance Model (Oprins, Burggraaff and

Van Weerdenburg, 2006; Oprins, 2008). It visualizes the complex cognitive processes of air traffic controllers (see Figure 1).

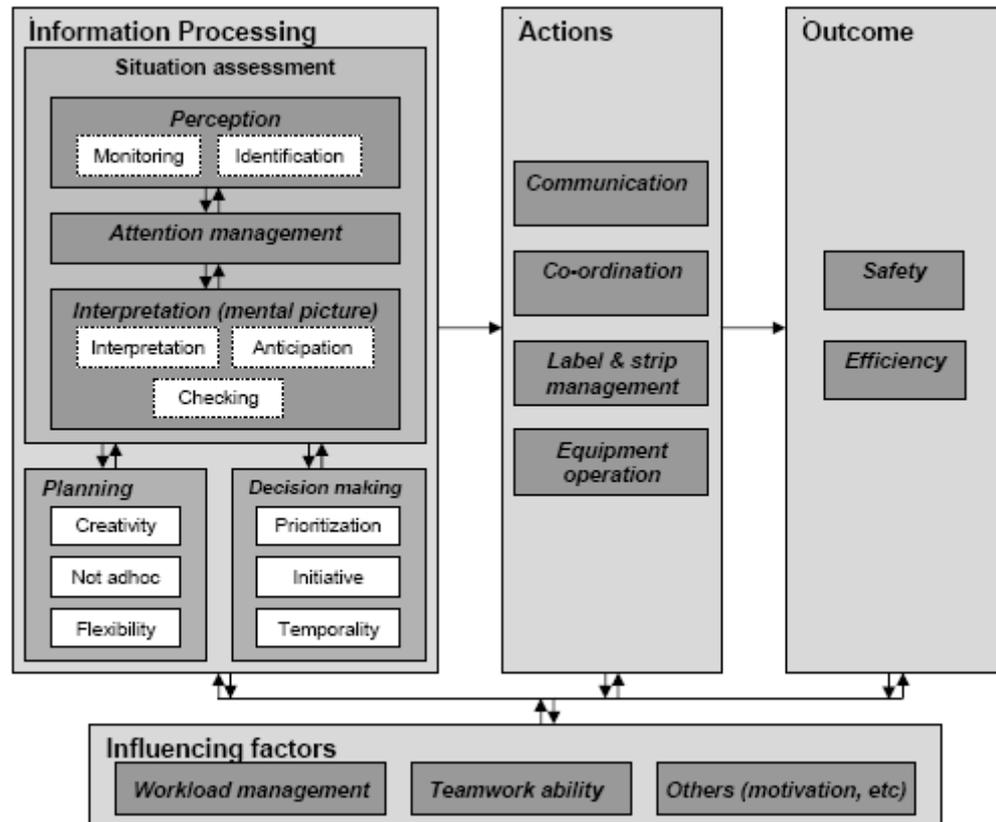


Figure 1. The ATC Performance Model (Oprins, Burggraaff & Weerdenburg, 2006)

The model shows the importance of cognitive processes. Information processing guides the actions which result in safe and efficient handling of traffic. This model has been applied as a general framework for selection and training design at LVNL. Since a few years, it is also used to assess the impact of developments in ATM system design on the human role of controllers in a paper study, called Human Factor Indication (HFI), and in real-time simulations.

Research on training performance of all trainees between 2003 and 2006, using the ATC Performance model, has shown that ineffective *situation assessment* and *workload management* are the two most important reasons for failing (Oprins, 2008). This suggests that these competences are more difficult to learn than others and require extra attention in designing less complex ATM systems.

2.2 Workload management

Controllers regularly switch between low and high mental workload, depending on the traffic situations (e.g., number of aircraft, complexity, peak periods). This is called workload management and differs between controllers (Averty, Collet, Dittmar, Athènes and Vernet-Maury, 2004). Controllers continuously apply strategies, which are individually different, to keep safety (e.g. conflict detection), efficiency (e.g. traffic delay) and their own mental workload ('personal efficiency') in optimal balance (Oprins, 2008). SA is needed to identify and enact a safe and efficient solution to solve specific (conflict) situations. In addition, controllers keep their own mental workload under control by adjusting their strategies towards less effortful if needed. If possible, they revert to routine actions, standard procedures and 'simple' solutions that need less attention and that gain time, for instance, by a reduction in radiotelephony (R/T). Depending on the evolving situation (routine – non-routine), they switch between low and high workload.

2.3 Situation awareness

A common assumption is that operators in dynamic and complex tasks such as ATC create a mental representation of the changing environment, which makes it possible to keep the relevant but transient information in working memory (Garland, Stein and Muller, 1999). Pattern recognition plays a central role; the controller groups aircraft in a certain way to memorize their positions. These patterns help them to create order in seemingly chaotic situations by streaming traffic flows. Much research has been done on how controllers develop the three-dimensional 'mental picture' of the traffic situation. This is usually referred to as *situation assessment*, defined as follows: 'The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future' (Endsley, 1988; 1995). *Situation Awareness* is considered the product of the process of situation assessment that takes place at three levels: perception (SA1), interpretation (SA2) and anticipation (SA3).

3 Automation in ATM

Automated ATM systems, as expected in the near future, should improve the controllers' performance by decreasing their workload and supporting their SA. This might reduce work complexity for controllers. With an expected increase in traffic in the future this seems to be a challenge. It is assumed that ATM is moving more towards monitoring or 'supervisory control' (SESAR, 2007; NextGen, 2007). Much research has been done on automation in ATM, for instance, in Free Flight concepts (RTCA, 1995, Endsley, 1997; Galster, Duley, Masalonis and

Parasuraman, 2001; Metzger, 2001; Metzger and Parasuraman, 2005). Three main issues are often evaluated in the design of automated systems: workload, SA and trust (Parasuraman, Sheridan and Wickens, 2008). These issues are considered very relevant in the function allocation between humans and systems.

3.1 Workload

Automation usually aims at a reduction of workload, (Metzger and Parasuraman, 2001; Metzger and Parasuraman, 2005). However, automation only offers benefits in workload reduction if the system is properly designed. Automation can produce both low and high extremes (Miller and Parasuraman, 2007). If automation is designed in a ‘clumsy’ manner, workload will be higher, e.g. if executing an automated function requires extra data entry or additional cognitive effort. Automation often converts manual tasks into monitoring tasks in which humans have a role as supervisor. This might impose considerable mental workload on operators even though they have to perform fewer (physical) actions (Metzger and Parasuraman, 2005; Miller and Parasuraman, 2007). For this reason, it is generally suggested to design an ATM system in which controllers maintain an active role in controlling the traffic (Metzger, 2001; Metzger and Parasuraman, 2005) especially when they are still responsible for safety. It is assumed that automating routine tasks will be beneficial for workload reduction, for instance, in display design for controllers (Metzger and Parasuraman, 2005).

3.2 Situation Awareness

A related issue that supports the active role of controllers refers to the ‘out-of-the-loop’ performance problem (Endsley and Kiris, 1995; Metzger, 2001). In case of automation failures system operators may have diminished ability to perform tasks manually, due to a reduced awareness of the state and processes of the system, i.e. SA. There are three reasons why this happens. First, monitoring tasks may lead to vigilance problems because controllers usually have much trust in the equipment. This decreases their alertness. Second, passive information processing seems to be inferior to active information processing in detecting the need for manual intervention and reorientation to the state of the system. Third, without any feedback, controllers are really out of the loop and they cannot assess the effectiveness of their requests and actions. Humans tend to be less aware of changes in the system when they are not in control. Moreover, reduced performance of manual tasks may also be the result of skill degradation due to the loss of SA (Miller and Parasuraman, 2007).

More automation can also increase SA (Endsley and Kiris, 1995). It is usually argued that automation should support SA by offering better and more integrated information to the controllers. Then they will be better able to distribute their attention, and SA will be improved

by a strong reduction of workload. A partial automation strategy should keep the negative and positive effects in balance.

3.3 Trust

Trust usually influences the human's dependence on automation. Human operators do not always use automated systems in the way that designers intended, because they trust their own abilities to control more than the system (Parasuraman and Wickens, 2008; Parasuraman, Sheridan and Wickens, 2008). An important factor is the reliability of automation. Trust in the system can also be too high. This may cause overreliance and failure to monitor the raw information sources that are input to automation, called 'complacency' (Parasuraman and Wickens, 2008). This can be avoided by subsequent exposure to automation's imperfection.

4 Future ATM developments

Increased automation is a central part of most ATM developments. It aims at increasing safety, efficiency and capacity of the future ATM systems while keeping work complexity acceptable. Automation is being considered for all phases of flight and for all involved parties. This means that both the role of the airline and air traffic control will be affected. Automation is prominent in both the European Union's Single European Sky ATM Research initiative (SESAR) and Federal Aviation Administration's NextGen program (SESAR, 2007; NextGen, 2007). These programs are aimed at solving the inefficiencies that exist in ATM today. The inefficiencies are largely a result of the fragmented development that comes from the fact that each nation is responsible for the efficient and safe management of its own airspace. It has resulted in each country setting up its own air traffic management with only the most necessary coordination with its neighbours. Though often well suited to the local situation it has resulted in a wide variety of solutions making the whole system unnecessarily complex and expensive for the airspace user. The phenomenon is seen world wide with Europe being one of the regions suffering the most from these negative consequences.

One important challenge for the ATM concepts, where automation may play an important role, is the need to find a good balance between the preferences of an individual flight and the need to organise efficient traffic streams. Also, the human factor impacts (workload, SA) must be taken into account.

4.1 Arrival Management

Arrival management is the area of ATM that deals with air traffic in the last phase of an arrival flight. It is concerned with the planning and controlling of aircraft that are landing at an airport. At most airports, the ATM system determines a landing sequence that is used by air traffic controllers to efficiently guide aircraft to the runway. It is one area where the different needs of individual flights (e.g., time schedules, preferred flight profiles for landing) become apparent and can lead to problems (e.g. delay). The SESAR Target Concept, that describes the future European ATM system, says the following on high density Terminal Manoeuvring Area (TMA) operation: ‘In high density traffic terminal areas (depending on the airport and/or the time), an efficient airspace organisation combined with advanced airborne and ground systems capabilities will be deployed to deliver the necessary capacity, maintain safe separation and minimise the environmental impact. The concept recognises that when traffic density is high the required capacity may only be achieved at the cost of some constraint on individual optimum trajectories’ (SESAR, 2007, p. 24). This definition allows for both ground and air-based systems to be developed to support the arrival management task.

4.2 Arrival management at Schiphol Airport

At Schiphol Airport, high traffic numbers and bunching associated with peaks puts high demands on the controllers’ competences and this makes work complex. Arrival traffic is fed from the Area of Responsibility (AoR) of Amsterdam Area Control Centre (ACC) into the Schiphol (TMA) via three entry points called Initial Approach Fixes (IAF). ACC controllers are required to deliver the arriving aircraft via the IAF within the target of plus or minus two minutes from the Expected Approach Time (EAT; i.e. planned time). Traffic streams in the TMA are subsequently merged, using radar vectoring, for the landing runway(s) in use. The LVNL arrival management strategy aims at reducing complexity for the controllers as well as increasing accuracy of delivery at the IAF, without compromising safety. There are three reasons for increasing the accuracy of traffic delivery.

First, merging traffic in the TMA is challenging because of the irregularity associated with the two minute margin over the IAF. This activity would be easier if all traffic would, on average, be delivered to the TMA closer relative to its planned time to increase predictability for the TMA controllers. Second, more accurate delivery is marked as an enabler for noise- and environmentally friendly conflict-free routes in the TMA. The current system cannot accommodate fixed, noise friendly, conflict-free routes in the TMA without a reduction in capacity. Third, improved accuracy translates to predictability and transparency to the airspace users. Pilots will be better able to manage the most efficient flight profile if their flight is planned well ahead, resulting in a reduction in delay and fuel burn. Therefore Amsterdam ACC

controllers need to deliver aircraft at a much higher accuracy to meet the target. In the LVNL ATM System Strategy a target is assumed of less than plus or minus 30 seconds.

To mitigate the potential increase in workload (Boudes and Cellier, 2000), system support is foreseen to enable the increased accuracy performance target to less than plus or minus 30 seconds. The Speed And Route Advisor tool was designed to support the controllers to meet this target.

5 SARA

5.1 Basic functioning

The SARA tool operates by providing controllers a speed and route combination for every inbound flight. The speed and/or route combination is displayed to the controllers and it will allow them to give a single speed and route clearance to the aircraft for the entire descent. A single clearance will have the potential advantage that it will decrease the workload for the controllers and aircrew. It will also allow the aircrew to optimally use the Flight Management Computer (FMC) in the descent, thereby optimizing the descent profile as much as possible within the active constraints.

The SARA logic works as follows. When a flight arrives within radar coverage of the ATM system the system plans its landing time. Subsequently, based on the landing time the Estimated Approach Time (EAT) is determined. The EAT is the time slot when the aircraft is planned to pass the IAF. The EAT at the IAF functions as an important metering point in arrival management. SARA uses the trajectory predictor function in the ATM system to predict when the aircraft will be at the IAF. This predicted time is called the Estimated Time Over (ETO). SARA then compares the difference between EAT and ETO, and SARA will calculate a new speed and route combination that will minimise the difference between the EAT and ETO if the difference is 30 seconds or larger. The speed and route combination is then presented to the controllers as an advice that will bring the aircraft at the IAF within 30 seconds of the planning. The advice is integrated in the aircraft label on the controllers' display, as shown in Figure 2.



Figure 2. A SARA advice as displayed in the aircraft's track label for the RTS

With SARA, controllers remain in control and are fully responsible for separation of the traffic (Prevot, Lee, Callantine and Smith, 2003). SARA only supports with calculating the speed and route combination best suited to meet the planning.

5.2 The impact of SARA on air traffic controllers

With SARA, the operation at Schiphol will gradually change from a tactical first-come-first-serve operation towards a time-based operation. These operations might have a quite large impact on the controllers' SA, and hence his subsequent capacity to stay in control (e.g. Metzger, 2001) However, the degree to which SA is affected depends on the specific operational design and task allocation between humans and systems (Miller and Parasuraman, 2007). The SARA tool could help controllers with instructing the right speeds and routes to aircraft in order to meet a specific waypoint on time. This might decrease their workload as once the instruction is given the controllers mostly need to monitor the follow up. Only in case of a conflict, they would need to give an updated instruction.

With SARA, the time-dimension will become more prominent in the controllers' mental picture in order to plan, prioritize and sequence flows, as well as to assure separation. This requires more anticipation and strategic thinking than nowadays. In their current way of working, their decisions are based on certain three-dimensional patterns of aircraft in which exact timing over a waypoint is a less crucial factor. Being in time at a waypoint within small margins changes the LVNL controllers SA because more 'thinking-in-time' is required than they are used to. Currently the controllers are more 'thinking-in-distance' and this determines how they sequence

the arrival traffic. Consequently, with SARA tactical control will move towards more strategic control (Oprins et al., 2009).

In addition, SARA implies that certain tasks of controllers are moved to the system. This might have impact on their workload and SA. Currently, controllers determine the speeds and routes for aircraft. SARA will help them in the decision making process by providing speed and route advisories. Controllers might lose their feeling of control when their work moves too much towards supervisory control (Galster et al., 2001). They might have difficulty to trust the system when solutions are in conflict with their own plan and their SA might be undermined (Parasuraman, et al. 2008). In other words, they cannot use their own strategies for traffic handling anymore. Dependent on the specific design of SARA, controllers could have less insight into the specific flight paths of aircraft. Consequently, it might make it difficult for them to update their SA if manual interventions are needed in case of system failures and other circumstances (e.g. weather) in which SARA may not work. Switching between these automated (routine) and manual (non-routine) operations can substantially increase their workload. It depends on the frequency of using conventional methods to which extent the controllers can act as the fallback.

In order to address these issues the SARA concept was tested in two settings, a Real-Time Simulation (RTS) and an Operational Trial. A RTS was chosen to evaluate the concept in a controlled experimental environment. Subsequently the Operational Trial was executed to further evaluate the concept in a real-life work environment. In both experiments, the influence of SARA on controllers' performance, workload and SA was investigated.

6 Study 1: Real Time Simulation

6.1 Experimental design

The Real Time Simulations were performed at NLR's ATC Research Simulator (NARSIM). The experiment was conducted during two days in which eight LVNL controllers participated (N=8). A single simulation run involved two controllers working in tandem for parts of the LVNL managed airspace (Amsterdam ACC sector 1 and sector 2) with two pseudo-pilots. The controllers and the pseudo-pilots communicated using R/T. Four identical runs were executed simultaneously.

Two familiarization runs were executed for each pair of controllers to familiarize themselves with the simulator the SARA Human Machine Interface (HMI) and the SARA way of working.

Next, the controller pairs executed four experimental runs. For comparison purposes, the same traffic sample was used for all runs. However, to avoid familiarization with the traffic sample, the aircraft call signs were shuffled between each run. Furthermore controllers switched working positions to also avoid effects resulting from the familiarity of the controllers with the traffic for a specific sector and inter-controller working strategies. The measured traffic sample contained 18 flights with destination Schiphol (EHAM). The four experimental runs consisted of two baseline runs and two SARA runs. Run 1 resembled current operations and functioned as a baseline in which controllers had standard system support and delivered aircraft at the IAF with an accuracy of plus or minus 120 seconds or less compared to the EAT. Run 2 functioned as a second baseline in which controllers had a stricter time target similar to the SARA runs (less than plus or minus 30 seconds) and limited system support. The support consisted of a delta time (ΔT ; EAT – ETO) presented in the aircraft label. In runs 3 and 4, SARA provided speed-only advisories, and speed and route combinations respectively. The properties of the simulation runs are depicted in Table 1.

Table 1. Properties of the simulation runs

Run	IAF target time (sec)	System support
1	Within +/- 120	Standard
2	Within +/- 30	Delta T in label
3	Within +/- 30	SARA speed
4	Within +/- 30	SARA speed & route

Quantitative and qualitative data was gathered during and after each simulation run. First, the accuracy with which the controllers managed to meet the EAT for each aircraft was measured. This measurement was called ‘EAT adherence’. As a subjective measure of workload the Instantaneous Self Assessment (ISA) was used. Controllers were prompted for input every three minutes. Objective measures of workload consisted of calculating the total number of R/T calls (i.e. radiotelephony; the verbal instruction administered to the aircrew), the average time spent on R/T by each controller, and the number of instructions entered into the system through the Touch Input Devices (TID; i.e. after instructions are given to the aircrew the controller enters them into the system via a TID). Directly after each simulator run, the controllers filled in an adapted version of the SASHA-Q Situation Awareness questionnaire (Dehn, 2008). Additionally, these questionnaires also contained open questions regarding workload, usability and acceptance. Interviews were held after each run to obtain in-depth information regarding their experiences with SARA. During the runs, four human factor observers were taking notes, one per controller pair.

6.2 Results

Repeated measures analyses of variance (ANOVA) were used for statistical comparisons. Partial eta-squared (η^2) is given as a measure of effect size. Pairwise comparisons were performed where appropriate with Bonferroni corrections. For each analysis an $\alpha < .05$ was used. Since the main objective of SARA was to reduce the variability of traffic delivery over the IAF this outcome is incorporated in the workload and SA graphs for comparison purposes.

6.2.1 EAT adherence

Data was obtained for 18 flights in the four experimental runs and was analyzed for missing values and outliers. Data was gathered for four pairs of controllers. The results showed a significant delivery accuracy improvement when SARA was used ($F(3,63) = 40.918, p < .001, \eta_p^2 = .661$; see Figure 3).

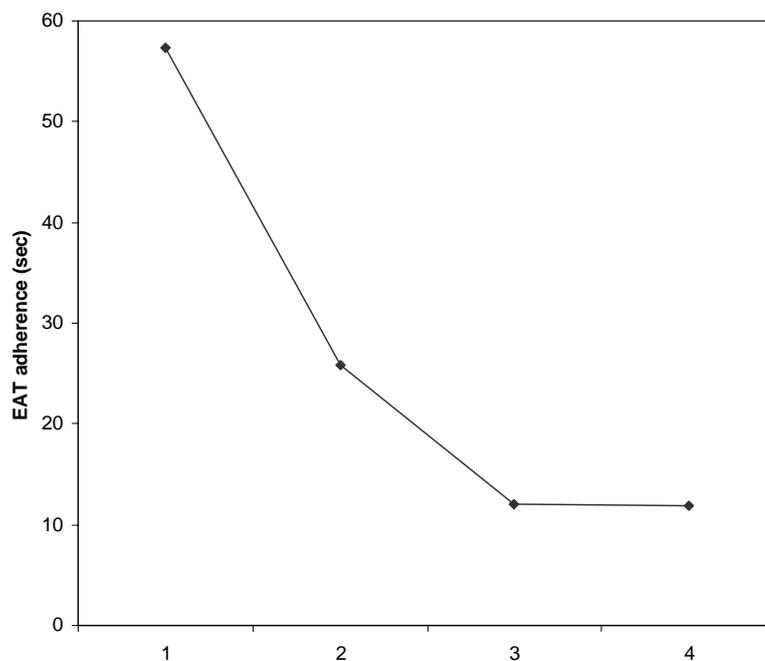


Figure 3. EAT adherence

The average absolute EAT adherence improved from the two baseline runs (run 1 and 2) to the two SARA runs (around 57 and 25 seconds accuracy to around 12 seconds accuracy; run 3 and 4). No significant differences were found between the speed only and the speed & route configurations (run 3 and 4). Interestingly, setting the target at less than 30 seconds and providing the controllers with limited system support (a delta T in the aircraft label; run 2) already significantly improved the accuracy to approximately 25 seconds.

6.2.2 Workload

6.2.2.1 ISA.

Eight LVNL controllers produced nine ISA scores each during each run (run 1 to 4). A significant effect was found between the four runs, $F(3,68) = 17.256, p < .001, \eta_p^2 = .432$. Workload in the SARA runs (run 3 and 4) was rated lower than the second baseline (run 2). Run 2 imposed a significantly higher workload on the controllers compared to the average of their ratings of the other runs, $p < .01$ (run 2 vs. run 1, 3 and 4). Run 4 (speed and route) was rated to be as equally demanding as run 3 (speed-only), $p = .701$. Compared to the EAT adherence results it seems that additional mental effort is required to achieve the improved accuracy. The results are depicted in Figure 4.

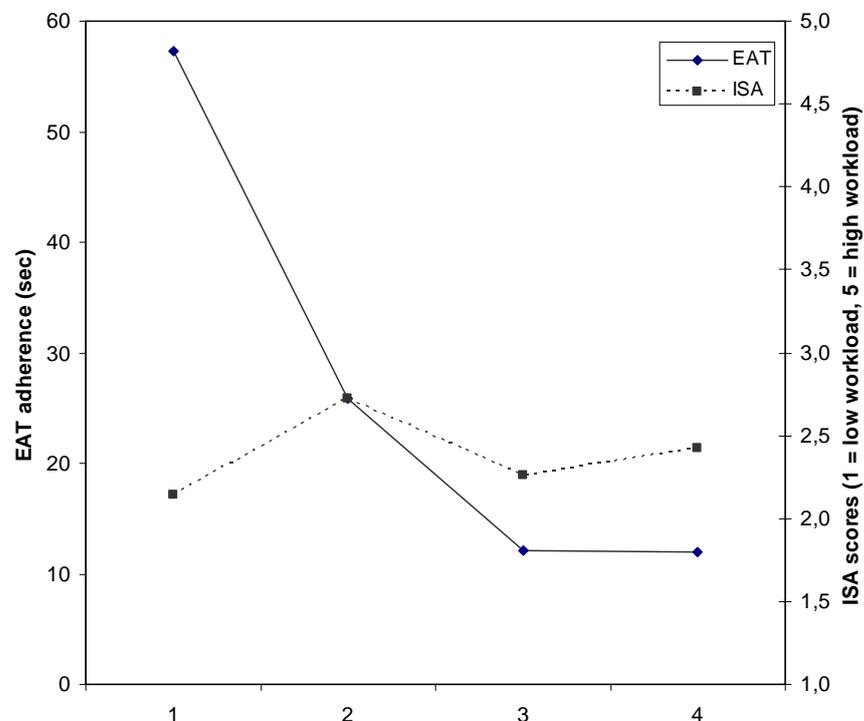


Figure 4. ISA scores

6.2.2.2 R/T calls.

After removing one outlier from the dataset seven measurements were obtained for the total number of R/T calls for eight LVNL controllers. A significant effect was found for this type of workload measure ($F(3,3) = 21.985, p < .05, \eta_p^2 = .956$). The SARA speed and routes run (run 4) required the lowest number of calls. The number of calls in this run was found to be less than baseline run 2 and the SARA speed-only run (run 3). A potential difference was found between run 4 and baseline run 1 ($p = .067$). SARA run 3 did not differ from the two baseline runs (run 1 and 2). The two baseline runs did not differ from each other. The results are depicted in Figure 5.

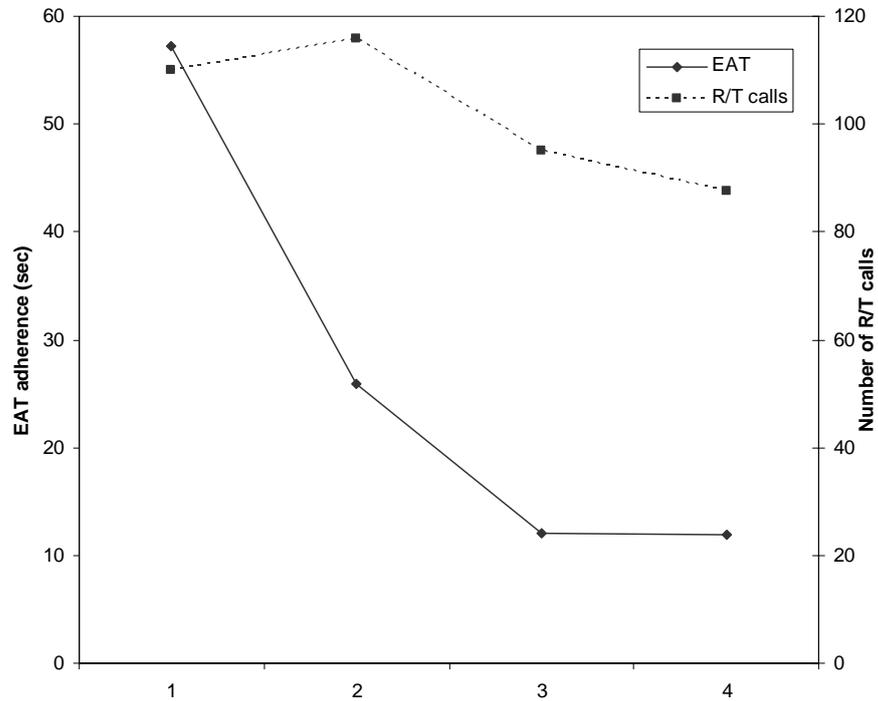


Figure 5. Number of R/T calls

6.2.2.3 R/T Time.

Eight measurements were obtained for the total time spent on R/T calls (in seconds) for the four simulation runs. An ANOVA showed significant differences between the four runs, $F(3,4) = 28.951, p < .01, \eta_p^2 = .956$. The lowest amount of time spent on R/T was found in the SARA speed & route run (run 4). There were no differences found between the first baseline (run 1) and the SARA speed only run (run 3). However, these two runs showed a reduced amount of R/T time compared to the second baseline (run 1 and 3 vs. run 2). No differences were found between the two baseline runs 1 and 2. See Figure 6 for the means for time spent on R/T.

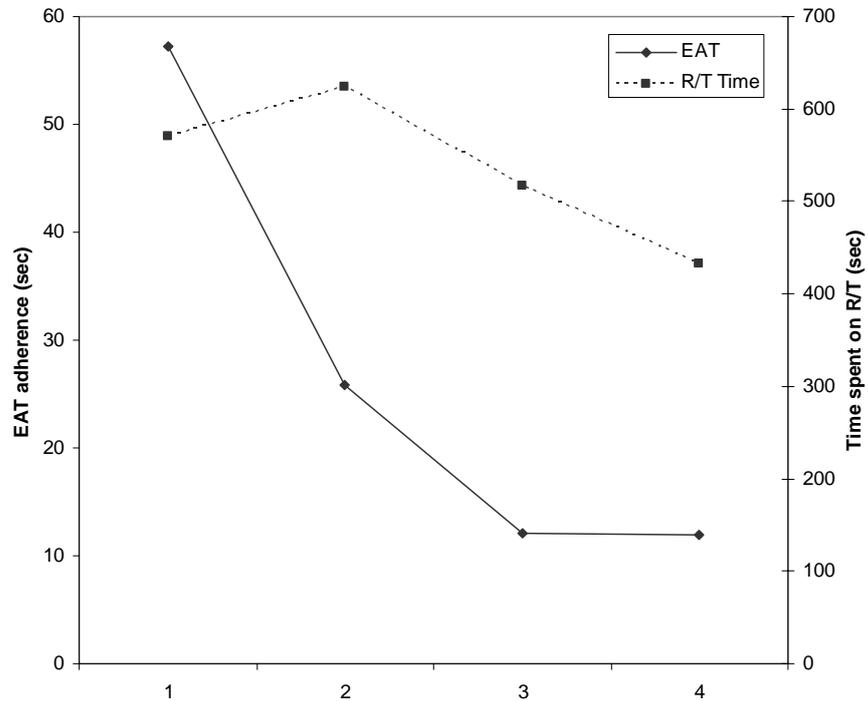


Figure 6. Time spent on R/T

6.2.2.4 TID inputs.

Eight measurements were obtained for the number of TID inputs for the four simulation runs. An ANOVA showed significant effects for the number of TID inputs, $F(3,4) = 11.091, p < .05, \eta^2 = .893$. The lowest number of inputs was found in the SARA speed & route run (run 4) compared to baseline run 2 and SARA run 3. A potential difference was visible between baseline run 1 and SARA run 4, $p = .051$. The highest number of inputs was found in baseline run 2 and potentially with baseline run 1, $p = .081$. The results are depicted in Figure 7.

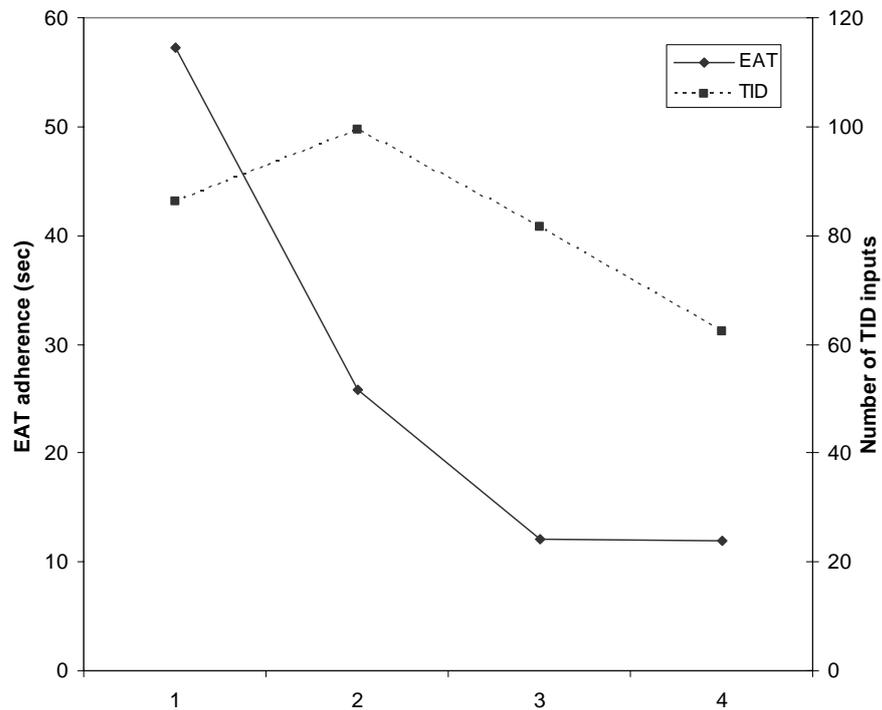


Figure 7. Number of TID inputs

6.2.3 Situation Awareness

The questions from the SASHA-Q questionnaire were averaged to serve as a total SA score for each controller (N=8). Four questions were used that were applicable to both the SARA runs (run 3 and 4) and the baseline runs (run 1 and 2). The Repeated Measures ANOVA showed a significant difference in SA scores between the four runs ($F(3,29) = 37.304, p < .001, \eta_p^2 = .794$). SARA runs 3 and 4 showed lower SA ratings compared to the two baseline runs 1 and 2. No significant differences were found between the two SARA runs (run 3 and 4) as well as between the two baseline runs (run 1 and 2). The results are depicted in Figure 8.

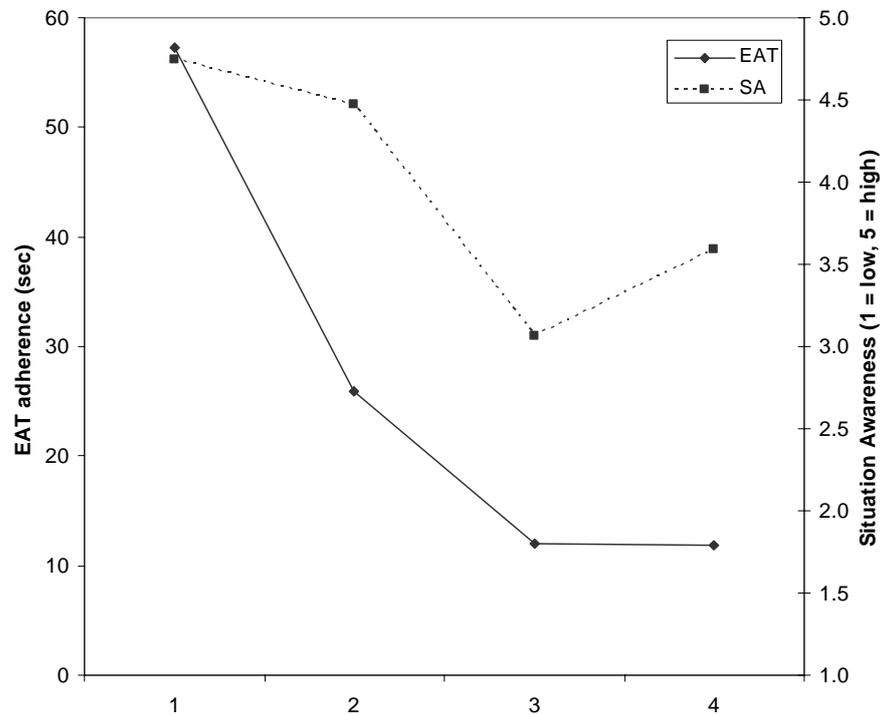


Figure 8. Situation Awareness

7 Study 2: Operational Trial

7.1 Experimental design

Ten controllers participated that were previously involved in the RTS or in earlier development of SARA. To minimize the impact of the trial on normal operations a speed-only SARA was evaluated. The trial was executed during eight three-hour periods (mornings and afternoons) in off-peak traffic which accumulated to 24 hours of measurement time. Baseline measurements were taken for comparison purposes on the same day as the trial period with comparable traffic and weather conditions.

To minimize the impact on the ATC system an HMI was chosen that was not integrated into the system. That is, the SARA system only received information from the operational system and was not in interaction with it. This meant that the controller had to read the SARA advisories from the HMI that was positioned adjacent to their radar screen.

Quantitative and qualitative information was gathered during the trials. First, the EAT adherence was calculated in the same manner as in the RTS. Second, TID inputs and a post-trial

questionnaire were used as workload measurements. Third, SA was measured directly after the trial using the SASHA-Q. Fourth, Human Factors experts were present during and after the trial to observe and take notes of specific behaviour and comments that controllers made. Fifth, interviews were held post-trial to address specific issues that may have had occurred during the trial and to discuss the general concept. However, due to operational constraints it was not possible to acquire all measurements during the baseline condition. Therefore, in this period only the objective data could be captured. The subjective data was only captured during the SARA trial periods.

7.2 Results

A total of 260 flights were handled during the measurement period versus 249 in the baseline. Traffic volumes varied between 1 and 23 flights per hour over the IAF with an average of 11.5. Weather did not have a significant effect on traffic handling.

7.2.1 EAT adherence

A significant difference was found in EAT adherence between the SARA periods and the baseline periods ($F(1,502) = 187.6, p < .001, \eta_p^2 = .272$). As is depicted in Figure 9 the mean accuracy of aircraft delivery as well as the variability in delivery of traffic is improved.

7.2.2 Workload

An increase was found in the average number of TID inputs per flight between SARA and the baseline ($F(1,1278) = 12.24, p < .001, \eta_p^2 = .009$).

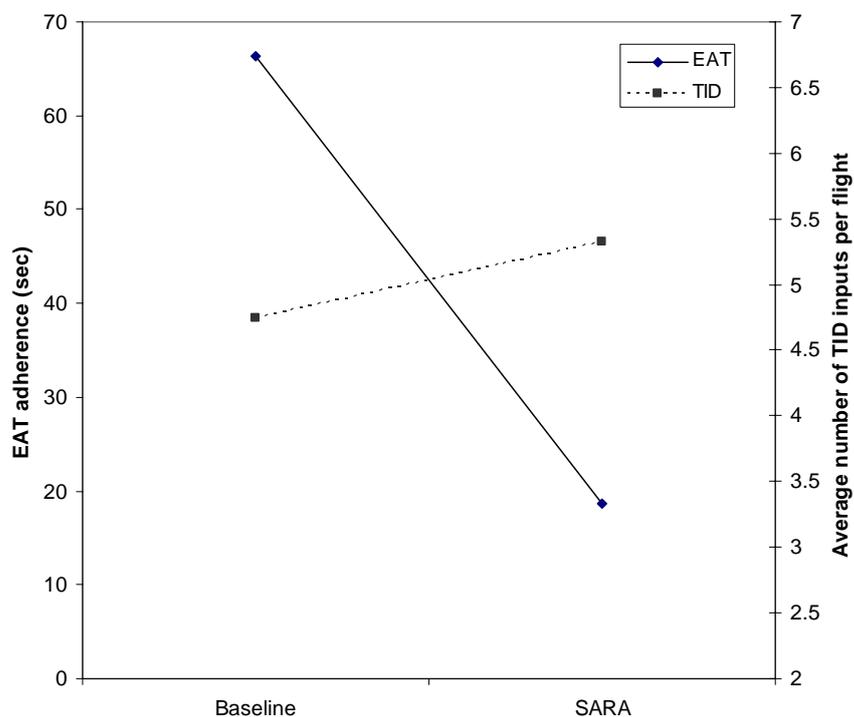


Figure 9. EAT adherence and TID inputs

More specifically it was found that the average number of speed instructions per flight increased from .54 in the baseline to .99 in the SARA condition ($F(1,1278) = 40.59, p < .001, \eta_p^2 = .031$). The average number of flight level instructions per aircraft increased from 2.11 to 2.28 ($F(1,1278) = 5.33, p < .05, \eta_p^2 = .004$). The average number of heading instructions per flight decreased from 1.76 in the baseline to 1.52 in the SARA runs ($F(1,1278) = 12.41, p < .001, \eta_p^2 = .010$). These results, although statistically significant, also show a very small effect size (smaller than .1) indicating a negligible increase in workload. These findings may have been inflated by the number of TID recordings. That is, small significance values in combination with very small effect sizes indicate a difference in TID inputs with little meaningful impact.

In the questionnaires the average workload value (3.31) was tested against a reference value (3) indicating the baseline. A significant increase in perceived workload by the controllers was found using a one-sample t-test ($t(26) = 2.264, p < .05$).

7.2.3 Situation Awareness

The SASHA questionnaire results showed an average SA score of 3.24 (out of a range of 1, low SA to 5, high SA; $SD = .55$). The minimum average SA score was 2.43 and a maximum average SA score of 4.38 was noted. These numbers could not be compared against a baseline value



because the questionnaire was not administered during this period due to operational constraints. Therefore, these numbers should be understood in relation to the controller's comments that were captured and analysed in the debriefing sessions.

It was mentioned that in periods of low traffic, few differences were noted in terms of SA. In periods of higher traffic numbers comments were made about the difference in traffic handling strategies between what SARA suggested and what controllers normally would do. This sometimes led to situations in which controllers accepted advisories, but were not completely sure how the situation would evolve. This made it more difficult for them to understand when to intervene when necessary. This sometimes resulted in relatively late deconfliction of traffic. With SARA controllers seemed to monitor the traffic for conflicts more than normally. For example, in one situation a controller was monitoring the two aircraft and was unaware that a third (trailing) aircraft was levelling off. As a result, this aircraft did not meet the altitude constraint at the IAF.

8 General Discussion

The aim of this study was to understand the impact of SARA on controllers' performance, i.e. their ability to more accurately deliver traffic at the IAF. Furthermore, the experiment aimed to understand the impact of this support tool on controllers' workload and SA.

8.1 Performance

The results showed that with the support of SARA controllers were able to deliver traffic more accurately. In the RTS an initial gain in delivery accuracy was seen even when controllers had minimal system support (only a delta T in the aircraft label; run 2). However, with the aid of SARA this accuracy was further improved (run 3 and 4). Furthermore, in the Operational Trial the delivery accuracy as well as variability of traffic delivery was improved when the controllers used SARA, even in real-world circumstances.

The stricter focus on time (30 seconds vs. 120 seconds) had a large influence on the working strategies of the controllers. Nowadays time is of lesser importance since controllers focus on creating 5 nm sequences. When doing so they most often meet the required time over the IAF. Therefore, in present operations time is of less importance. With a target of within plus or minus 30 seconds controllers will have to invest more effort to meet the target and requires a different mind set by the controllers. This may mean that generating sequences of 5 nm may not be enough to meet the target, but that more precise actions are required. In SESAR and NextGen

there is a large focus on stricter time-based operations in which the entire trajectory of an aircraft is planned from gate to gate together with strict fixed times over waypoints (SESAR, 2007). The experiences in this study may shed some light on the expected future working methods for controllers.

To fully support controllers in meeting their performance target a proper design of the tool is essential. One design decision that was proven to be successful was the use of instant feedback. Instant feedback enhances the controllers' ability to stay on top of the traffic situation by 'scanning the traffic, identifying the need for an action, and issuing a proactive instruction' (Prevot et al., 2003, p. 8). When controllers issued instructions to the aircraft controllers immediately perceived the consequences of their instructions through the change in delta T for that aircraft on their HMI. This is in contrast to current operations in which, due to technical reasons, the change in delta T only slowly changes closer to the IAF. Any delay in such feedback may cause the controllers to become behind in handling other traffic with a reduced performance and potential reactive controlling behaviour as a result. The behaviour of the delta T function with SARA speed and/or route was shown to be a successful implementation of instant system feedback.

8.2 Workload

With SARA subjective workload (ISA scores) did not increase compared to the baseline in the RTS. Objective workload (number of R/T calls and TID inputs) reduced compared to the baseline, especially for SARA with speed and route options. However, in the Operational Trial these results were not replicated. That is a significant but negligible increase in TID inputs was recorded, especially the number of speed instructions. Metzger and Parasuraman (2006) showed that communication and coordination tasks can be a considerable source of workload to controllers, especially under high-traffic conditions. Our results therefore seem to indicate that controllers' workload is likely to remain equal if not reduced when using SARA. The RTS results also showed that workload (subjectively as well as objectively measured) was highest for the run with minimal system support (run 2). Some controllers however, mentioned that they felt that, with SARA active, other activities needed to be performed and not necessarily more or less. That is, a change in working method was experienced by some controllers that may have resulted in a lower physical workload, but a similar mental workload. This can partly be explained by the fact that the controllers were somewhat unfamiliar with SARA. Although they had previous experience, they were still learning, which may have increased their mental workload. The fact that objective workload was reduced in the RTS points at an expectable decrease of workload in the future, especially under normal traffic conditions. However, specific comments were made about an increase in the monitoring of traffic compared to the normal operation that resulted in a



reduction in available time for managing other traffic. This seemed only to occur in high traffic density situations. In low traffic density conditions, controllers reported a decrease in their workload.

With the current implementation of SARA only arriving aircraft were provided with speed and/or route advisories. This led some controllers to mention that potential difficulty of working with two working methods: the arrival traffic stream under SARA advisories and the departure traffic stream under 'normal' control which could potentially add to their workload. It is important to address these issues in terms of training or procedures if SARA is to be used in this way.

8.3 Situation Awareness

Interestingly, in the RTS SA was highest in the two baseline runs (run 1 and 2) and dropped significantly when SARA was used (although still rated as above average). The ATC Performance Model suggests that SA is one of the prime information processing components of controllers (Oprins et al., 2006; Oprins, 2008). Controllers build up a mental picture by perceiving, interpreting and anticipating on the traffic stream. Based on this continuous process controllers decides on the required instructions for aircraft in the traffic stream. With SARA part of this activity is performed by the automation since it provides controllers with advisories that have not been part of their mental processes. Alternatively, the mental picture created by the controllers and the resulting instructions may differ from the solutions provided by SARA. Therefore, it is understandable that controllers rated their SA as lower compared to the baseline scenarios. Controllers mentioned that with SARA they felt 'less engaged' in the traffic situation compared to the baseline runs. This refers to the out-of-the-loop-performance problem as described by Endsley and Kiris (1995) in which operators have a reduced awareness of the state and processes of the system. SARA had taken over some tasks of controllers and therefore they felt less 'in control'. It was mentioned that with SARA they followed an advice and monitored its progress. Some controllers mentioned that they felt that, because SARA produces an advice at the FIR entry to meet the time over the IAF, this would mean a solution for more than meeting the time alone, i.e. a conflict-free advice. This sometimes lead to controllers solve conflicts late rather than early. Furthermore controllers mentioned that when SARA was active, they felt that they spent time to understand SARA's 'plan' as part of an effort to create a mental picture of the traffic situation in contrast to generating their own plan.

Controllers changed their interaction with the SARA tool during the course of the simulations. It was observed that controllers regained some of their SA by not adhering to advisories all the time. It was observed that sometimes advisories were used as a 'general guidance' to give an

aircraft a speed that would more or less be adequate to meet the time over the IAF. When SARA provided subsequent advisories, these would first be evaluated by the controllers for their usefulness before they were instructed. This was contrary to the controllers' behaviour early in the RTS where every given advice was accepted and instructed by them. This suggests that sufficient familiarization with SARA is required before it can be implemented and used in an operational setting. In the RTS controllers became more used to SARA after a few hours. The experience during the preparation of the Operational Trial showed that little time was required to familiarize themselves with the system and the new way of working. Because all controllers were previously engaged in the RTS or earlier developments of SARA, they used the tool as a guidance mechanism straight away. This has provided useful information for training purposes when implementing SARA in an operational setting.

8.4 Future directions: system support in Time-Based Operations

In this study two versions of automation support were tested: a speed-only options and a speed and route option. A third version is foreseen that will incorporate Conflict Management (CM) to provide controllers with conflict-free speed and/or route advisories. This particular version of SARA was out of the scope of our evaluations and it was therefore not possible to investigate its consequences for controllers. However, as shown previously, performance and workload under mature Free Flight may hint at considerations for design and implementation of this particular version (Galster, Duley, Masalonis and Parasuraman, 2001). In the present study a decrease in SA was found with the use of automation support with speed and/or route options. A further impact on SA is expected when advisories are conflict free. Checking advisories for potential conflicts becomes unnecessary since these are calculated to be conflict free. With controllers partially 'out-of-the-loop' (Endsley and Kiris, 1995), they may not be up to the challenge, due to complacency issues, in case conflicts are not resolvable by SARA (Galster et al., 2001; Wickens, Mavor, Parasuraman and McGee, 1998).

Several studies have provided recommendations on the out-of-the-loop-performance problem. It has been argued that to keep controllers in the loop, they should retain some of their responsibilities and automation should support them in their decision-making (Wickens et al., 1998). Previous research also showed that controllers more easily accept automation support if they are in command (Prevot et al., 2003). That is, they found that solutions over which controllers had a choice were more readily accepted compared to solutions that appeared automatically. For automation support tools, and in this particular case SARA, this may mean that controllers should be able to have a choice about which speed and/or route combination they want to issue rather than having the automation present a single solution only. One option for keeping controllers in the loop may be to present various conflict-free solutions which the

controllers can choose from, possibly with visual support. This way, controllers remain in control of the traffic, whilst the automation is able to support controllers in their decision-making.

8.5 Conclusion

This study demonstrates that decision support tools can improve performance whilst reducing workload and supporting Situation Awareness. SARA was designed for air traffic controllers to achieve a more accurate traffic delivery. The findings from the RTS and the Operational Trial indicate that system support, such as SARA, is needed to improve the performance of the current day operations without increasing controllers' workload. A form of automation support was designed that provides controllers with an advice on the optimal solution. Since SARA takes over some tasks of controllers, SA was affected in a certain way. Controllers had the feeling of being less engaged in their tasks (cf. out-of-the-loop problem; Endsley & Kiris, 1995). Their working methods changed more than expected, moving towards time-based operations. However, it was noticed that more familiarity with SARA improved the controllers' SA and decreased their workload. In the future, SARA will be further developed in such a way that controllers are still in-the-loop to avoid a possible loss of SA, for instance, by offering various solutions from which the controllers can choose. The specific design of the tool and the controllers' familiarity with it determine the degree to which a sufficient SA could be maintained.

SESAR and NextGen are aiming for stricter time-based-operations with subsequent automation support for controllers. More automation is needed to answer the challenges facing the aviation industry in terms of safety, efficiency and environment. However, in the current system, and likely in the future, humans play a central role (Parasuraman and Wickens, 2008). Therefore the responsibilities between the humans and the systems should be carefully assessed when designing and implementing automation solutions.

9 Acknowledgements

This project is sponsored by the Knowledge and Development Centre (KDC); a cooperation of Dutch ANSPs, airlines, airports and research centres. We would like to thank the SARA project team (LVNL, NLR, Eurocontrol, KLM and Boeing) for their efforts into the development of SARA and their support in this project.

10 References

- Averty, P., Collet, C., Dittmar, A., Athènes, S. & Vernet-Maury, E., (2004). Mental workload in air traffic control: an index constructed from field tests. *Aviation, space, and environmental medicine*, 75(4), 333-341.
- Boudes, N. & Cellier, J., (2000). Accuracy of estimations made by air traffic controllers. *The international journal of aviation psychology*, 10(2), 207-225.
- Dehn, D. M., (2008). Assessing the impact of automation on the air traffic controllers: the SHAPE questionnaires. *Air traffic control quarterly*, 16(2), 127-146.
- Endsley, M.R., (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society 32nd annual meeting*, 97-101.
- Endsley, M.R., (1995). Towards a theory of situation awareness in dynamic systems. *Human factors*, 37(1), 65-84.
- Endsley, M.R., (1997). Situation awareness, automation and free flight. *Proceedings of the 1st Air Traffic Management R&D seminar Saclay, France*.
- Endsley, M.R., & Kiris, E.O., (1995). The out-of-the-loop performance problem and level of control in automation. *Human factors*, 37(2), 381-394.
- Galster, S.M., Duley, J.A., Masalonis, A.J. & Parasuraman, R., (2001). Air traffic controller performance and workload under mature Free Flight: conflict detection and resolution of aircraft self-separation. *The international journal of aviation psychology*, 11(1), 71-93.
- Garland, D.J., Stein, E.s. & Muller, J.K., (1999). Air traffic controller memory: capabilities, limitations and volatility. In: Garland, D.J., Wise, J.A. & Hopkin, V.D. (Eds.). *Handbook of aviation human factors*, (pp. 455-496). Mahwah, NJ; Erlbaum.
- Merwe, G.K. van de, Oprins, E., Plaat, A. van der & Eriksson, F., (2009). The influence of automation support on air traffic controller behaviour with a Speed and Route Advisor function. *Proceedings of the International Ergonomics Association Conference, Beijing*.
- Metzger, U., (2001). The role of air traffic controller in future Air Traffic Management: an empirical study of active control versus passive control. *Human factors*, 43(4), 519-528.
- Metzger, U. & Parasuraman, R., (2005). Automation in future Air Traffic Management: effects of decision aid reliability on controller performance and mental workload. *Human factors*, 47(1), 35-49.

- Metzger, U. & Parasuraman, R., (2006). Effects of automated conflict cuing and traffic density on air traffic controller performance and visual attention in a datalink environment. *The international journal of aviation psychology*, 16(4), 343-362.
- Miller, C.A. & Parasuraman, R., (2007). Designing for flexible interaction between humans and automation: delegation interfaces for supervisory control. *Human Factors*, 49(1), 57-75.
- NextGen (2007). *Concept of operations for the next generation air transport system*. Version 2.0.
- Oprins, E., (2008). *Design of a competence-based assessment system for ATC training*. Doctoral dissertation, Maastricht University.
- Oprins, E., Burggraaff, E. & Weerdenburg, H. van, (2006). Design of a competence-based assessment system for ATC training. *The international journal of aviation psychology*, 16(3), 297-320.
- Oprins, E., Zwaaf, D., Eriksson, E., Merwe, G.K. van de, Roe, R., (2009). Impact of future time-based operations on Situation Awareness of air traffic controllers. *Proceedings of the 8th Air Traffic Management R&D seminar Napa, California*.
- Parasuraman, R., Sheridan, T.B. & Wickens, C.D., (2008). Situation awareness, mental workload, and trust in automation: viable, empirically supported cognitive engineering constructs. *Journal of cognitive engineering and decision making*, 2(2), 140-160.
- Parasuraman, R., & Wickens, C.D., (2008). Humans: still vital after all these years of automation. *Human factors*, 50(3), 511-520.
- Prevot, T., Lee, P., Callantine, T. & Smith, N., (2003) Trajectory-oriented time-based arrival operations: results and recommendations. *Proceedings of the 5th Air Traffic Management R&D Seminar Budapest, Hungary*.
- RTCA (1995). *Report of the RTCA Board of Director's Select Committee on free flight*. Washington, DC: Author.
- SESAR (2007). Deliverable 3: The ATM Target Concept (DLM-0612-001-02-00a). SESAR consortium.
- Schneider, W., (1990). Training high-performance skills: fallacies and guidelines. In: Venturino, M. (Ed.). *Selected readings in human factors* (pp. 297-311). Santa Monica, CA: The human factors society.
- Wickens, C.D., Mavor, A., Parasuraman, R., & McGee, J., (1998). *The future of air traffic control: human operator and automation*. Washington, DC: National Academy Press.