Executive summary

Impact of future time-based operations on situation awareness of air traffic controllers

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. 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 executed at Air Traffic Control the Netherlands (LVNL).

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.

Description of work

A ground-based Speed and Route Advisor (SARA) tool has been designed to help Area Control (ACC) controllers with achieving a higher punctuality. A future follow-up for SARA could be an air-ground agreed Controlled Time of Arrival (CTA).

Results and conclusions

The SARA real-time experiment results showed that this tool definitely decreases 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 more focusing 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, increasing automation (cf. SARA advisories) could be in conflict with the controllers' own plan of traffic handling. They could loose a certain 'feeling of control' and ultimately their SA. This refers to the 'out-of-the-loop' problem of automation. 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 resulting set of SA requirements on the design of such controller support tools shall be addressed in future developments of time-based operations in ATM.

Impact of future time-based operations on situation awareness of air traffic controllers

Nationaal Lucht- en Ruimtevaartlaboratorium, National Aerospace Laboratory NLR

Anthony Fokkerweg 2, 1059 CM Amsterdam,
P.O. Box 90502, 1006 BM Amsterdam, The Netherlands
Telephone +31 20 511 31 13, Fax +31 20 511 32 10, Web site: www.nlr.nl

NLR-TP-2009-236

Impact of future time-based operations on situation awareness of air traffic controllers

E.A.P.B. Oprins[1], D. Zwaaf[1], D.J.F. Eriksson[1], G.K. van de Merwe and R. Roe[2]

[1] Luchtverkeersleiding Nederland

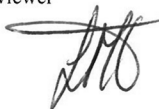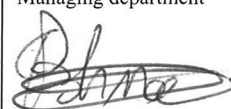[2] Universiteit van Maastricht

This report is based on a presentation held the Air Traffic Management R&D Seminar, Napa Valley (CA), U.S.A., 29th June 2009.

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

Approved by:

Author	Reviewer	Managing department

Contents

This page is intentionally left blank.

1 Introduction

Future air traffic management (ATM) systems are expected to migrate towards trajectory-based operations as proposed in the Single European Sky ATM Research (SESAR) and NextGen concepts. The target is a planned operation mutually agreed between stakeholders such as airlines, Air Navigation Service Providers (ANSPs) and airports. Trajectory-based operations implicitly include extra focus on the dimension of time. An operation that requires meeting a 3D waypoint at a specific time calls for a different approach to handling traffic with increasing automation. Today, air traffic controllers use tactical speed, route, altitude and vector instructions, based on a first-come-first serve principle. The inclusion of time may have a large impact on the controller's Situation Awareness (SA) because of the focus on time instead of on distance. Previous research has shown that increasing automation of ATM systems could decrease controllers' SA, due to vigilance decrements, availability of information and other factors [1] [2]. These human factor impacts must be addressed and evaluated when designing new ATM systems and support tools, especially in areas characterized by great complexity and high-density such as Schiphol Airport.

This paper explores how the controllers' SA is affected by time-based operations. It emphasizes the reduction of task complexity and creates a basis for a gradual implementation. First, a short literature review on SA in ATM is provided. Second, two operational concepts of time-based operations addressed at LVNL are explained. This involves the Speed and Route Advisor (SARA) tool and the concept of Controlled Time of Arrival (CTA). Third, some empirical results of real-time simulations (RTS) with SARA are presented. Finally, a set of SA requirements on the design of controller support tools in future time-based operations in ATM are derived from the findings in the literature and from the preliminary results of the SARA experiment.

2 Situation awareness in ATM

2.1 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 [3]. Pattern recognition plays a central role; the controller groups aircraft in a certain way to memorize their positions. These patterns help them to order a seemingly chaos 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' [4]. Situation awareness (SA) is considered the product of the process situation assessment that takes place at three levels: perception (SA1), interpretation (SA2) and anticipation (SA3). Attention management strategies are crucial to keep this ever changing 'picture' up-to-date [5]. Controllers continuously apply strategies, which are individually different, to keep safety (conflict detection), efficiency (traffic delay) and their own mental workload ('personal efficiency') in optimal balance [6]. SA is needed to identify and enact the most 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. This is called *workload management*. If possible, they revert to routine actions, standard procedures and 'simple' solutions that need less attention and that gain time, for instance, by a lower load of radiotelephony. Depending on the evolving situation (routine – non-routine), they switch between low and high workload (cf. vigilance).

2.2 The competence SA in ATM

Internally, LVNL is coping with a shortage of controllers, which is not uncommon at many busy and complex ATC units. Due to the complex cognitive nature of the ATC task only a small number of people is able to acquire the required competences within a reasonable period of training [7]. LVNL is trying to solve this problem by improving selection and training, and also by designing new ATM systems that reduce task complexity. As a starting point, a competence analysis was performed at LVNL based on literature research and workshops with controllers. This has resulted in the ATC Performance Model [8] [9], see Figure 1.
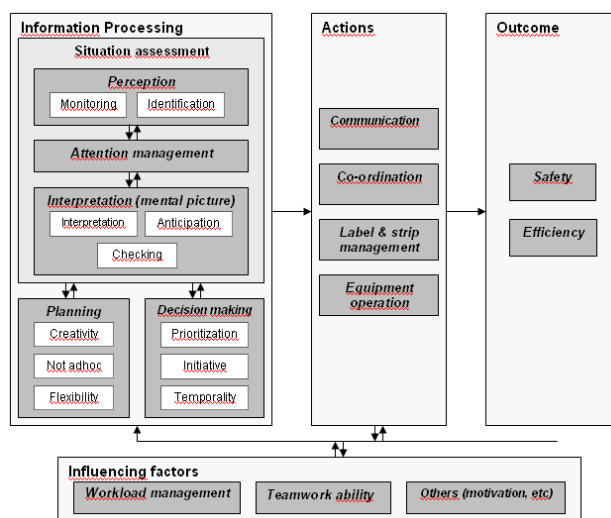

Figure 1. The ATC Performance Model [8] [9]

The model shows the importance of cognitive processes in which *situation assessment* plays a central role. Information processing guides the actions and this results in safe and efficient traffic handling. One important influencing factor is *workload management*. Research on training performance of all trainees between 2003 and 2006, based on this model, has shown that ineffective SA and workload management are the two most important reasons for failing [9]. This suggests that these competences are more difficult to develop than others and require extra attention in designing less complex ATM systems. Within this context, setting human requirements on SA and workload management serve as an input in the design of ATM systems at LVNL. In addition, the impact on SA, workload management, and other competences, as part of the model, is systematically assessed. This is done for each change, by a paper study and by measuring them in real-time simulation (RTS) experiments.

2.3 Automation effects on SA

Previous research has shown that increasing automation as expected in future ATM systems could reduce task complexity, and could have an effect on SA in various ways. A possible risk of more automation is often referred to as the 'out-of-the-loop' performance problem [1]. In case of automation failures system operators may have diminished the ability to perform tasks manually, due to reduced awareness of the status and processes of the system (SA). There are three reasons why this happens. First, monitoring tasks may lead to vigilance problems. Alertness decreases and controllers usually have much trust in the equipment. Second, passive information processing seems to be inferior to active information processing when detecting the need for manual intervention and reorientation to the state of the system. Third, people are really out of the loop without any feedback, and they cannot assess the effectiveness of their requests and actions. On the other hand, more automation can also mitigate a reduction in SA [1]. In a more monitoring role, controllers are better able to spread their attention, especially when the system provides superior, integrated information to the controllers. In addition, a reduction in SA may be mitigated by a strong reduction of workload. A partial automation strategy should keep the negative and positive effects in balance. It is usually argued that routine tasks should be fully automated to reduce workload, while automation should support SA by offering better and more integrated information to the controllers.

These issues have been addressed in research on ATM system design [2], and are mentioned by SESAR [10] and NextGen [11]. ATM is also moving towards more monitoring (cf. 'supervisory control'). Human-centred design in ATM suggests that routine tasks such as radiotelephony should be automated (cf. datalink), that information should better be displayed to controllers for supporting SA, and that decision support tools are needed to choose the right solutions. However, ATM system designers are still searching for the right balance in automation, also in relation to fallback systems (machine or human) [10]. Clear and more detailed guidelines for

human-centred ATM design are not available yet. Often, new ATM technologies are the starting point from which consequences for the human role are derived, while human requirements should drive the design as well.

3 ATM strategy of LVNL

3.1 Automation drivers

Concepts of new ATM systems include various requirements to increase the level of automation. For LVNL, many of these requirements have been imposed by the market, and national- and European legislation. The home carrier and other airlines are requiring an ever improving cost-effectiveness, punctuality and capacity, the government is demanding a safe and noise friendly operation and the European Union is enforcing the SESAR concept [10]. These requirements cannot be met with the current level of automation. Part of the LVNL strategic development is to improve the predictability and accuracy of arrival traffic for Amsterdam Airport Schiphol [12]. It is the largest airport in the Netherlands, handling about 430.000 flights per year.

In addition, expedited by the existing shortage on controllers, reduction of work complexity has become a central principle in the LVNL strategy. This implies that more candidates would be able to become a competent controller and that the transition to new ATM systems for current operational personnel would be sufficiently easy. More automation can help with making the task easier for controllers.

3.2 ATM strategy

Based on these automation drivers, LVNL has developed an ATM strategy that meets these requirements. An important enabler is the introduction of CDAs for arriving aircraft. To achieve this, LVNL will first need to develop new procedures and systems to improve its Arrival Management (AMAN) process, called inbound planning. These procedures include an improved planning on predefined arrival routes, based on Precision Area Navigation (P-RNAV) and better weather information in the Terminal Area (TMA). The use of tactical vectoring will reduce if system support improves the execution of the inbound planning. For controllers, it will be more difficult to tactically adjust the arrival flow in order to maximise runway throughput. The arrival flow will need to be fixed and optimized at an earlier stage, which could lead to a reduced runway throughput. However, LVNL studies on P-RNAV routes and procedures in the Schiphol TMA indicate that this can be mitigated by a more accurate traffic delivery at the IAF. If aircraft arrive within an accuracy of less than plus or minus 30 seconds of their expected approach time (EAT) at IAF instead of plus or minus 120 seconds being required nowadays, controllers will be

able to maintain a runway throughput of approximately 35 aircraft per runway (per hour). Experiments have been conducted using accurate traffic samples of realistic operations. In conclusion, developing an improved planning alone will not be sufficient to meet the runway throughput requirements. It will only work when aircraft will meet the planning. In other words, the ATM strategy of LVNL implies a shift towards time-based operations in accordance with the operational concepts of NextGen and SESAR.

3.3 Time-based operations in ATM

The heart of a new envisaged ATM operation as described in SESAR [10] and NextGen [11] is the concept of Trajectory Based Operations. The objective is to optimize the use of airspace to cope with the increasing number of flights and to keep the ATM network affordable. The ATM system should be a performance-based, service-oriented operational concept. Stakeholders should improve their planning and meet their respective target times to enable a smooth operational process. The reference business trajectory should be a contract to what the Airspace User agrees to fly and the ANSP and Airport agree to facilitate. Good planning and execution of that planning, *a time-based operation*, will be key to the success of the operational concept. For the ANSP, a time-based operation will pose new requirements on the architecture of the ATM system and air traffic control procedures. The function of the controller is to maintain and expedite a safe and orderly flow of traffic. Improved accuracy in meeting a planning as produced by an AMAN system becomes a hard requirement.

3.4 Support tools in time-based operations

To improve punctuality both pilots and controllers will need automated assistance. The support tools that LVNL is investigating, vary from a fully ground-based trajectory prediction solution, the SARA tool, to an air-ground negotiated CTA at a given fix, which allows an optimal FMS determined route and flight parameters.

SARA was developed to support area controllers in delivering aircraft at IAF with high precision (cf. less than plus or minus 30 seconds instead of plus or minus 120 seconds or less). SARA generates a speed and route advise that will allow the controller to give a single clearance to the aircraft for the whole descent. A single clearance will decrease some aspects of the workload for the controller and aircrew. It will also allow the aircrew to use the FMS in the descent to optimise the descent profile. The main benefit of SARA, being a ground-based tool, is that controllers can use it for all arriving flights. Specific aircraft equipment is not needed. Therefore, LVNL believes that SARA will be a useful tool for controllers that can be implemented reasonably soon. CTAs, as part of the SESAR concept [10], could subsequently be implemented. Implementation of CTAs into the operational ATM system requires Required Time of Arrival (RTA) technology in the FMS system of aircraft. The challenge lies in the fact

that various aircraft have different FMS systems on board. Some FMS do not have RTA build in yet, and even if they do, the operation could differ. In the CASSIS project [13], in which LVNL participates, some early applications of CTA have been described and tested in fast-time simulations. The graduate transition from various varying SARA concepts to SARA designs and towards deployment of full CTAs is assumed extremely important from human factors perspective with SA being a central issue.

3.5 SA in future time-based operations

For the controller, the operation will gradually change from a tactical first-come-first-serve operation towards a more strategic time-based operation. These operations might have a large impact on the controller's SA, and hence the capacity to act. However, the degree to which SA is affected depends on the specific operational design and task allocation between humans and systems. The ground-based SARA tool could help controllers to instruct the correct speed and route to aircraft in order to arrive at a fix on time. This might decrease the workload as, once the instruction is given, the controller only needs to monitor the execution of the instruction. The controller would only be required to give an instruction in case of a conflict

The CTA concept takes the monitoring role a step further. The ground and airspace user agree a time to meet the fix. The FMS will determine the right flight profile to meet the objective; thus, in CTA the controller appears to be more out-of-the-loop. It also depends on the specific, more detailed design properties of SARA or CTA. Some specific changes of SA and possible risks are expected, which should be considered in the design and must be verified in real-time simulations.

In both SARA and CTA, controllers will have to incorporate time as a fourth dimension in their 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 the current way of working, the controllers mentally form three-dimensional patterns of aircraft on a certain moment of time, on which they base their decisions. Being in time at a waypoint with high accuracy changes the controller's 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. Tactical control will move towards more strategic control with a larger planning horizon. In the CTA concept, the controllers will provide pilots with time constraints. This is new for both because nowadays the clearances only include speed, altitude and heading. As stated in ATM research on human factors in SESAR [10], more long-term planning could mean that certain responsibilities will move from the radar controller to the planner or even to the pilot in managing traffic flows in radar control.

In addition, both SARA and CTA imply that certain tasks of controllers are moved to the system. Currently, controllers determine the speeds and routes for aircraft. SARA will support

them with achieving higher accuracy by providing speed and route advisories. In case of CTA, the FMS determines how flights will achieve the required time of arrival in terms of speeds and altitudes. Controllers might loose their feeling of control when their work moves too much towards supervisory control. They might experience difficulties with trusting the system when solutions are conflicting with their own plan and their SA might be (adversely) undermined. In other words, they cannot use their own strategies for traffic handling anymore. Dependent on the specific application of SARA and CTA, controllers could have less insight into the specific flight path of aircraft. This will definitely decrease their SA. Consequently, it might make it difficult for them to regain their SA if manual interventions are needed in case of system failures and other circumstances (e.g., weather) in which SARA does not work. Switching between these automated (routine) and manual (non-routine) operations can increase their workload substantially. It depends on the frequency of using conventional methods to which extent the controllers can be the fallback.

Especially during transition towards the new concept, the new system might not work optimally. Circumstances such as traffic mix (e.g., RTA-equipped and non-equipped aircraft in CTA operations) or non-nominal situations (e.g., bad weather, runway changes, traffic bunches) might force the controllers to switch between conventional and new operational concepts. When such a switch is required, their mental picture will have to immediately change. This poses a critical demand on cognitive flexibility, raises the controller's workload, and carries the risk of reduced control performance. As a result, the chance for human errors may increase. Therefore, this switching must be limited as much as possible and the transition to the new situation should be made gradual. In addition, in case of CTA, it should be clear for controllers which aircraft are RTA-equipped and which are not. This enhances their SA because they can easily incorporate this information into their mental picture.

4 Real-time simulation Experiment

LVNL has conducted real-time simulation (RTS) experiments with SARA but not yet with CTA. Therefore, the focus here is on the preliminary results of the RTS experiment, which is executed for two design alternatives for SARA. The expected impact on the controller SA was assessed, as well as the relationship between workload and SA based on the ATC Performance Model [8] [9] (see Figure 1).

4.1 Design of SARA

The SARA tool relies on several functions in the ATM system: Inbound Traffic Planner (or AMAN), Surveillance Data, and a Trajectory Predictor (TP). The performance of these support

functions determines the performance of SARA. SARA processes a flight in seven steps.
These steps are as follows:

1. The flight is activated in the ATM system and becomes part of the AMAN planning.
2. Once the planning is considered stable, SARA starts to function.
3. In SARA, the Expected Approach Time (EAT) is used for the flight as input.
4. The TP is used to determine the current position of flights and planning (surveillance data). The TP is also used to calculate the flights Estimated Time Over (ETO) the IAF. For this calculation, it is assumed that the route entered into the ATM system will be the route flown.
5. The EAT is compared with the ETO. If the difference is outside a set bandwidth (plus or minus less than 30 seconds at IAF), the process of generating advisories will be initiated. These advisories can consist of speed (concept 1) or a combination of speed and route (concept 2).
6. An iterative process is started where the TP is used to calculate a speed or a speed and route combination that will bring the aircraft to the IAF such that the EAT and ETO is below the threshold value.
7. Once a solution within the bandwidth is found, it is communicated to the controller. The SARA advice will be integrated in the HMI. Figure 2 illustrates how speed advisories are displayed in the aircraft label, and EAT and delta Time (T; EAT-ETO) in the stacklist.

Figure 2. The SARA HMI.

In this process, the controller remains fully responsible for separating the traffic. SARA only calculates speed (concept 1) or speed and route combination (concept 2) to execute the planning in the most optimal way. However, SARA does not provide conflict free solutions with both concept 1 and 2. This implies that controllers must deviate from advisories when they evoke

conflicts. In future developments of SARA or other time-based operations, a conflict detection and resolution step could be added to the process. In this RTS experiment, concept 1 and concept 2 have been evaluated providing SARA advisories to LVNL controllers only, or both to LVNL and Maastricht UAC (MUAC) controllers.

4.2 Methods

The SARA RTS was performed at NLR's ATC Research SIMulator (NARSIM). The experiment was conducted during four days. At the first two days, eight LVNL controllers participated (N=8), while at the other two days four LVNL controllers and four MUAC controllers participated (N=2x4). Because of the length of this paper and the uniformity of the interpretation of the data, only the results from the first two days, with LVNL controllers, will be presented.

A single simulation run involved two controllers and two pseudo-pilots working in tandem for Amsterdam Area Control (ACC) sector 1 and sector 2. The pseudo-pilots had radio contact with the controller for the specific sector. Four identical runs were executed at the same time. Each pair of controllers started with two training runs to get familiarized with the simulator and the SARA HMI. The results do not include the training runs. Next, the pairs executed four experimental runs. For comparison purposes, the same traffic sample was used for the four runs. The measured traffic sample contained 18 arrival flights. Between each run, the aircraft callsigns were shuffled and controllers switched working positions. Therefore, effects resulting from the familiarity of the controller with the traffic sample, the controllers' familiarity with the traffic for a specific sector and inter-controller working strategies were minimized.

The four experimental runs consisted of two baseline runs and two SARA runs. Run 1 mimics 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 (T; EAT – ETO) presented in the aircraft label. In runs 3 and 4, SARA provided respectively speed-only advisories, and speed and route combinations. Table 1 gives an overview:

Table 1. Properties per run

Run	Target time over IAF	System support
1	Within plus or minus 120 sec	Standard
2	Within plus or minus 30 sec	Delta T in label
3	Within plus or minus 30 sec	SARA speed
4	Within plus or minus 30 sec	SARA speed & route

During and after each simulation run, quantitative and qualitative data was gathered. The accuracy with which the controllers managed to meet the planned time over the IAF (EAT) for each aircraft was called 'EAT adherence'. During the runs, the Instantaneous Self Assessment (ISA) was used as a subjective measure of workload and prompted the controller for their input every three minutes. Furthermore, workload was measured objectively by calculating the total number of R/T calls, 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). Directly after each simulator run, the eight controllers were asked for their opinions on SA using an adapted version of the SASHA-Q questionnaire [14]. These questionnaires also contained open questions. Interviews were held after each run serving as a debriefing. The focus was on their experiences with working with SARA and its acceptability. During the runs, human factor observers were taking notes. The questionnaires, interviews, and observations provided additional qualitative results.

4.3 Results

Repeated Measures Analysis of Variance (ANOVA) was used to compare the various runs with each other. Partial eta-squared (η_p^2) is given as a measure of effect size. Pairwise comparisons with Bonferroni corrections were performed where appropriate to calculate specific outcomes. For each analysis, an $\alpha < .05$ was used.

<u>EAT adherence</u>
The obtained data for 18 flights in the four experimental runs, handled by four controller pairs, were analyzed for missing values and outliers. The results showed a significant delivery accuracy improvement when SARA was used, $F(3,63) = 40.918$, $p < .001$, $\eta_p^2 = .661$. Figure 3 presents the means and standard deviations for each run (it should be noticed that the graphs in all figures are not trend lines but the points are interconnected as a visual aid).
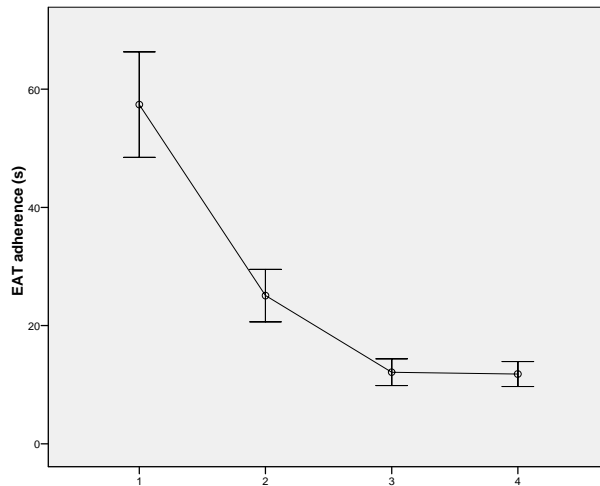
Figure 3. EAT adherence

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

Workload

For each run, nine ISA measurements were obtained from each of the eight controllers. An ANOVA on the ISA scores for the four runs showed a significant effect, $F(3,68) = 17.256$, $p < .001$, $\eta_p^2 = .432$. Workload in run 3 and run 4 only differs much with 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 4 (speed and route) was rated to be as equally demanding as run 3 (speed-only), $p = .701$. Figure 4 presents the means and standard deviations for each run for the ISA measures:
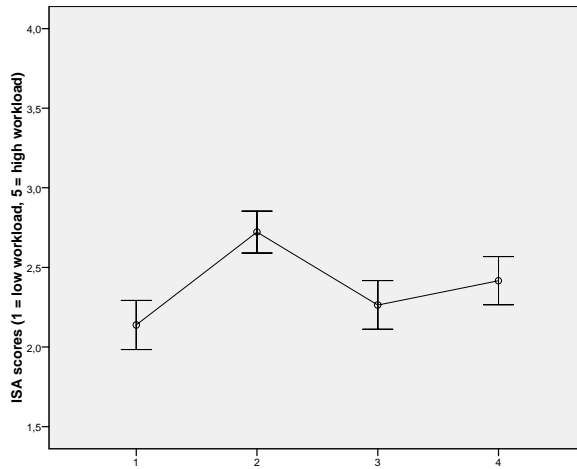
Figure 4. ISA measures

Seven measurements were obtained for the total number of R/T calls for the four simulation runs in two sectors (one outlier was deleted). These workload measures showed a significant effect in the ANOVA, $F(3,3) = 21.985$, $p < .05$ $\eta_p^2 = .956$. Run 4 (speed and routes) required the lowest number of calls; less than baseline run 2 and SARA run 3, and potentially with baseline run 1 ($p = .067$). SARA run 3 did not differ from the two baseline runs. The two baseline runs did not differ from each other. Figure 5 presents the means and standard deviations per run for the number of R/T calls:
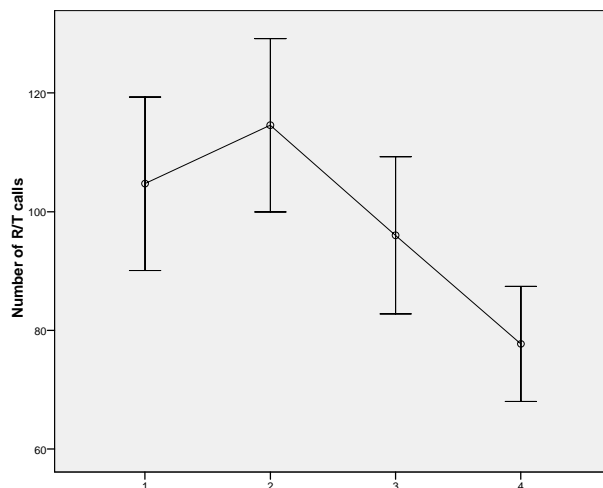

Figure 5. Number of R/T calls

Eight measurements were obtained for the total time spent on R/T calls (in seconds) for the four simulation runs in two sectors. Significant results were found in the analysis on the time spent on R/T, $F(3,4) = 28.951$, $p < .01$, $\eta_p^2 = .956$. The speed and route combination of SARA (run 4) resulted in the lowest amount of time spent on communication with the aircraft compared to the

other runs. The speed-only variant (run 3) resulted in less R/T than baseline run 2, but did not differ from the baseline run 1. No differences were found between the two baseline runs 1 and 2. See Figure 6 for the means and standard deviations for time spent on R/T:
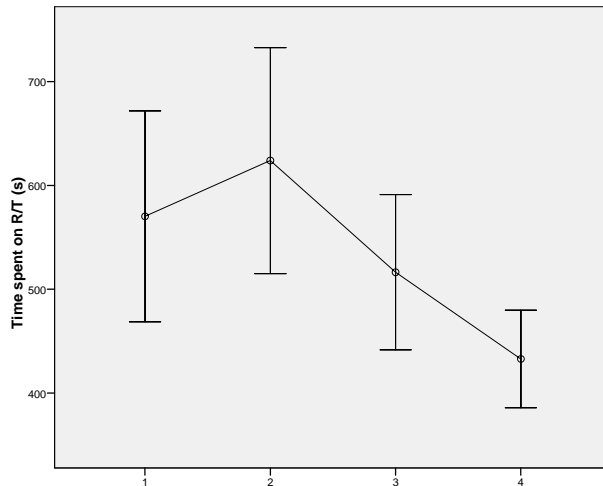

Figure 6. Time spent on R/T

Eight measurements were obtained for the number of TID inputs for the four simulation runs in two sectors. Significant effects were found for the number of TID inputs, $F(3,4) = 11.091$, $p < .05$, $\eta_p^2 = .893$. The total number of instructions was the lowest for SARA run 4 compared to baseline run 2 and SARA run 3. A trend was visible between baseline run 1 and SARA run 4, $p = .051$. Baseline run 2 showed the highest number of inputs compared to the SARA runs and potentially with baseline run 1, $p = .081$. Figure 7 presents the means and standard deviations for the number of TID inputs:
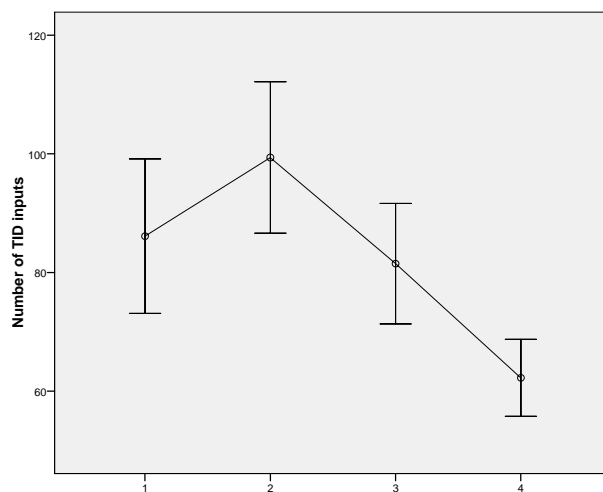

Figure 7. Number of TID inputs

Situation Awareness

Questions from the adjusted 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 and the baseline runs. A significant reduction in SA was found in an ANOVA $F(3,29) = 37.304$, $p < .001$, $\eta_p^2 = .794$. SARA runs 3 and 4 showed lower significance ratings compared to the two baseline runs 1 and 2. No significant differences were found between the two SARA runs as well as between the two baseline runs. Figure 8 presents the means and standard deviations of the SA ratings:
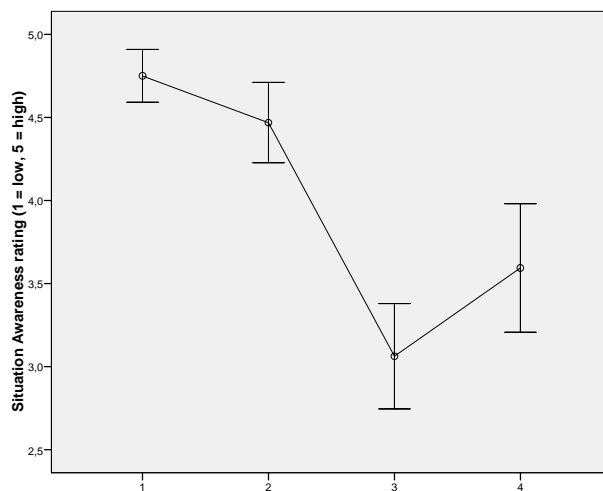

Figure 8. Situation Awareness

Qualitative results: the changed role of controllers

In general, the qualitative evaluation results based on the open questions, observations and interviews indicated that the eight controllers were quite positive about the possibilities that SARA offers. The tool makes it possible to be more accurate at meeting the EAT at the IAF. This would be much more difficult without SARA. However, the human role of the controller is assumed to change when SARA is implemented. With respect to SA, the findings pointed at two major changes.

First, a higher accuracy performance target for aircraft to be at the IAF (plus or minus less than 30 seconds instead of 120 seconds) was experienced as quite different, both with and without SARA. The observers noticed that controllers were sequencing arriving aircraft and solving conflicts differently in order to achieve that aircraft would be on time. For instance, extra vectors were given and level separation was applied instead of lateral separation. The sequence of aircraft was often different from what the controllers would normally create. This changes controllers' SA substantially because the time constraint gets more priority in the working

strategies. Controllers are used to creating solutions by use of more distance, rather than time. This explains why the controllers rated SA in the SARA runs lower than in the baseline runs as illustrated in Figure 7. Some controllers suggested that an introduction of time-based operations should start with forcing controllers in achieving more accuracy at IAF. As system support, only delta T in the label is needed and not necessarily SARA advisories. This might be useful for achieving a gradual transition towards the complete operational concept.

Second, automation in the form of speed (and route) advisories was experienced as a novelty. This automation takes over some mental processes from the controllers. The advisories were sometimes different from the controllers' own plan. For instance, speeds were proposed that resulted in a sequence that controllers never would choose, and some speed advises were 'unnatural' such as speed changes just before IAF and speed changes of 5 knots. Controllers found it difficult to follow advisories that deviate from their own preferred strategy. Their own mental picture changes when advisories are in conflict with this, which could decrease their SA. Controllers said that they did not feel in complete control of the traffic and that extra effort was required to maintain SA. The controllers were forced to check all advisories and ignored some due to conflicts or other reasons. Some of them noticed that it is so natural to accept advisories while they should not trust SARA in all cases (e.g. conflicts). For instance, controllers remained in a reactive mode by waiting for the first advisory before issuing the first instructions to the aircraft when entering the FIR. This is in contrast with normal operations to solve conflicts as soon as possible, preferably directly at the FIR entry. With respect to SA, there was a difference between run 3 (only speed advisories) and run 4 (speed and route advisories). The controllers argued that while extra vectors were required to absorb delay in run 3, published route options in run 4 enabled a more predictable traffic flow that enhances SA. In addition, there were individual differences between controllers. Some of them recognized SARA advisories as guidance rather than a compulsory instruction.

The controllers argued that this novelty increased their workload in the first SARA exercises. This explains the increasing values of the ISA measures presented in Figure 4. Controllers said that they spent much time on understanding what the logic of the system wanted, and they were often confused by the advisories of SARA which were contradicting compared to their own strategies. Hence, their physical activities did not change much, as supported by the objective workload measures presented in Figure 5 – 7, but the mental activities changed and increased. However, the controllers indicated also that they became used to the algorithm behind SARA. They started to predict which advisories SARA would produce after a few experimental sessions. This implies that there was a strong learning effect over time. More experience will most likely also increase their SA. In addition, the controllers argued that it might be problematic to use two different working methods. Currently, SARA is designed to handle only arriving traffic for Schiphol, but not for arriving traffic for other airports or for departure traffic.

5 Discussion and conclusions

In conclusion, the RTS experiment has shown that with system support of SARA a higher accuracy in meeting the EAT at the IAF is feasible. Even with minimal system support (delta T in aircraft label), the controllers were able to meet the target, but with SARA, further improvement in accuracy in meeting the EAT could be achieved. Differences in workload, both subjective and objective, were found between the various runs. The SARA runs (3 and 4) showed a decrease in physical activities (R/T calls, R/T time and TID inputs) compared to the baseline runs (1 and 2), but subjective workload was experienced as higher. The objective workload was lower in run 4 (speeds and routes) than in run 3 (speeds only). This is logical, because the SARA advisories reduce the actions required by the controller. SA was rated higher for run 4 than for run 3, because the published route options in combination with speed advisories resulted in a more predictable traffic flow. In run 3, with only speed advisories, controllers were forced to use unnatural vectors for achieving accuracy. However, this is probably also influenced by a learning effect. Given the fact that the experiment took only a few days, the chance that their workload would decrease after more experience in line with the objective workload measures is very high.

Changes in SA can explain why subjective workload is experienced as higher in the SARA runs than in the baseline runs, although the objectively measured workload indicates the opposite. As expected from the literature, there seems to be two major influences on SA. First, more focus on time (higher accuracy at IAF) instead of on distance changes the controller's three-dimensional mental picture. This results in a different way of sequencing arriving traffic and solving conflicts. Planning becomes more important. Second, increased automation by SARA advisories decreases the controller's SA, especially controllers with minor experience. The system replaces some mental activities of the controller, that is, considering appropriate speeds (and routes). Consequently, the controllers cannot always apply their own preferred strategy in the traffic handling if they accept the advisories. These strategies are assumed crucial for current controllers to be able to deal with changing traffic situations and workload [6]. It might decrease human flexibility to switch between various situations (routine vs. non-routine) and therefore it requires support tools such as SARA. The findings of the RTS also provide some proof for the theory of the 'out-of-the-loop' performance problem in increased automation as described [1]. The controllers experienced less awareness of the situation, less feeling of control, and sometimes they trusted the system too much. However, in the SARA concepts evaluated in the RTS, the controller is still fully responsible for ignoring or accepting SARA advisories and for the separation between aircraft. In other words, the controller is still completely 'in-the-loop' and makes the decisions, even though SARA calculates advisable speeds (and routes) to support the controller. The problem of vigilance caused by too much monitoring, risking that the

controller cannot regain SA while he functions as the fallback, is not applicable here. Additionally, SA is expected to increase when controllers gain experience and can predict SARA advisories. This definitely will decrease their subjective workload. The only problem might be that controllers get used to the advisories and cannot make the calculations anymore when SARA cannot be applied due system failures, weather circumstances, special traffic situations etc. The 'out-of-the-loop' problem could become important in further developed time-based operations when routes are dynamic, conflict management is automated and datalink is used. The controllers' SA will decrease because there is less need to form a mental picture if the system manages the traffic flows. They have minor insight into the planned flight path of aircraft. In that case, the human cannot be the fallback anymore in case of system failure. Finally, it must be noted that this SARA RTS experiment has certain limitations. First, the experiment was restricted to a few days with involvement of only a small number of participants. Definitive conclusions about the changed human role of controllers in terms of SA and workload cannot be given yet. Second, only two SARA concepts were evaluated with just one HMI. Alternative SARA designs can lead to other results. For instance, the experience of decreasing SA might change if the information is displayed in another way. Third, measuring learning effects, relevant for a possibly gradual implementation of SARA, requires a longer period of evaluation. Therefore, the preliminary results of this experiment should be verified in a broader context of literature research and related experiments (e.g. in SESAR).

5.1 SA requirements in time-based operations

A set of SA requirements for the design of controller support tools such as SARA and in CTA in future time-based operations can be derived from the findings in the literature and from the preliminary results of the SARA RTS experiment. It is important that such requirements guide the design instead of only deriving human factor consequences when the tools have been technically designed already. Starting point is that the support tools for controllers must reduce the ATM system complexity, as stated in the LVNL ATM strategy, next to other stakeholder requirements for cost-effectiveness, capacity, punctuality, safety and noise reduction. Based on the results described in this paper a following set of ten SA requirements for time-based operations in ATM could be:

- Achieving more accuracy in meeting the EAT requires a clear planning display for controllers that predicts time effects on their actions (cf. delta T) in order to support their SA.

- System support must take over complex calculations from the controllers (e.g. appropriate speeds) for keeping the workload and SA of an acceptable level when achieving more accuracy with the same traffic capacity.

- System support (e.g. speed advisories) must be so natural (cf. human-centered) that they can become part of the controllers' strategies for maximizing their SA.

- As long as the controller is responsible for the separation between aircraft, certain flexibility in using the system is needed (e.g. excluding single or multiple flights from SARA advisories) to let the controllers follow their own strategies based on SA.

- For constantly maintaining SA, the controller must remain active by taking his own decisions in the aircraft's flight path, eventually helped by system support.

- The right amount of information must be displayed in such a way that the controllers can form one integrated mental picture for maintaining their SA.

- If the system has become a conflict manager that has taken over the controllers' SA, the fallback cannot be the human anymore; the system must be fully reliable.

- The transition towards time-based operations must be done in small steps (system support, procedures) to become familiarized with differences in SA.

- A gradual implementation of time-based operations may not result in a mixture of working methods that increase the controllers' workload and decrease SA.

- In mixed traffic situations, the system must display which aircraft are equipped (e.g., with RTA) and which are not for supporting the controllers' SA.

Many SA requirements are strongly related to safety and might be addressed in a safety case, for instance, how reliable a fallback system must be or when the controllers´ workload is not acceptable anymore.

5.2 Final conclusion and future directions

In conclusion, time-based operations as planned in future ATM systems require well-designed support tools for controllers and a graduate implementation to compensate for the impact on controllers' SA and workload. At LVNL, the SA requirements are addressed in the further development of SARA and CTA operations in the coming years. An operational trial with SARA (only speed advisories) is planned in 2009 to further develop the Trajectory Predictor of SARA and to evaluate the impact on controllers in the operational environment. LVNL also participates in the project CASSIS that investigates CTA operations within the context of the SESAR concept [13]. Real-time simulations and operational trials for Schiphol Airport are planned in this project for the next two years. All these activities must contribute to a successful implementation of time-based operations at Schiphol Airport.

References

[1] 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.

[2] Endsley, M.R. (1997). *Situation awareness, automation and free flight.* Paper presented at the FAA/Eurocontrol Air Traffic Management R&D seminar. Saclay, France.

[3] 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.

[4] Endsley, M. (1995). Towards a theory of situational awareness in dynamic systems. *Human factors, 37 (1),* 32-64.

[5] Shebilske, W.L., Goetll, B.P. & Garland, D.J. (2000). Situation awareness, automation and training. In: Endsley, M.R. & Garland, D.J. (Eds.). *Situation awareness: analysis and measurement* (pp. 303-323). Mahwah, NJ: Lawrence Erlbaum Associates.

[6] Oprins, E. & Burggraaff, E. (2006). *Learnability of Situation Awareness and Workload Management in ATC training.* Proceedings of the 16th World Conference on Ergonomics.

[7] 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.

[8] 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.

[9] Oprins, E. (2008). *Design of a competence-based assessment system for ATC training.* Doctoral dissertation, Maastricht University.

[10] SESAR Deliverable 3 (2007). *The ATM Target Concept* (DLM-0612-001-02-00a). SESAR consortium.

[11] NextGen CONOPS (2007). *Concept of Operations for the Next Generation Air Transport System.* Version 2.0.

[12] Mijatovic, D., & Dijkstra, F. (2008). *Towards trajectory-based operations in arrival management at Schiphol Airport.* Submitted to the ATM R&D seminar 2009.

[13] CASSIS (2008). *CTA/ATM system integration studies (CASSIS): concept of operations* (D-LFV 2008-056339). CASSIS consortium.

[14] Dehn, D. M. (2008). Assessing the impact of automation on the air traffic controller : the SHAPE questionnaires. *Air Traffic Control Quaterly, 16(2),* 127-146.

Acknowledgments

Author Biography

Esther Oprins studied Dutch science and educational psychology at Utrecht University. She started as consultant and researcher at the Centre for Innovation of Training (CINOP) in Den Bosch, the Netherlands. She has been working as expert at Air Traffic Control the Netherlands (LVNL) for eight years now. During that period, she obtained her PhD at Maastricht University on the subject of assessment on competences of air traffic controllers. The focus in her research activities at LVNL have moved from selection and training towards ATM strategy development and human factors in ATM system design.

David Zwaaf studied Aerospace Engineering at Delft University of Technology. He started his career as a Systems Engineer at Airbus in Bristol, UK and continued later at LogicaCMG. For three years now he has been working as a Concepts Expert at Air Traffic Control the Netherlands (LVNL). In this role he expanded his System Engineering competences and focussed on ATM strategy and new ATM concepts such as a possible introduction of CTA applications.

Fredrik Eriksson studied microbiology at the University of Edinburg, UK. This led him to work with research in medical microbiology and immunology at the Wilhelmina Kinderziekenhuis in Utrecht., the Netherlands. He later worked with cold chain management and air transport of temperature sensitive pharmaceutical products. His current work is as concept expert at the ATM Strategy Development group at LVNL.

Koen van de Merwe has a background in cognitive psychology at the Leiden University, the Netherlands. He has been teacher of statistics for psychology students at Leiden University. Currently he is a researcher at the National Aerospace Laboratory NLR in Amsterdam, the Netherlands. His work focuses on the functioning of humans in cognitive complex environments. His research field includes civil and military aviation systems and has a strong focus on Air Traffic Control.

Robert Roe has worked as researcher, teacher and consultant in work & organizational psychology at the universities of Amsterdam, Delft, Tilburg and Nijmegen, the Netherlands. Morover, he has been director of the Work and Organization Research Center (WORC) in Tilburg and the Aeromedical Institute (AMI) in Soesterberg. He is currently professor of organizational theory and organizational behaviour at Maastricht University. His publications include books, chapters and articles on human resources management, automation, human factors, work performance and the study of dynamic processes in organizations. He is involved in projects in the domain of air traffic control.