

34th Conference of the European Association for Aviation Psychology
**Air Traffic Controller Competence Retention and Retention
Modelling: a preliminary study**

Petra ten Hove*, Guido Tillema and Jennifer Eaglestone

Royal Netherlands Aerospace Centre (NLR), Anthony Fokkerweg 2, Amsterdam 1059 CM, the Netherlands

Abstract

Air Traffic Controller (ATCO) error can have a huge impact on flight safety, therefore preventing skill decay is essential. Training can be a good remedy, but training is expensive and sometimes unnecessary. Understanding how and when skill decay occurs is essential in personalising retention training. This preliminary study examines methods for measuring (military) ATCO skill decay. Experimental sessions were conducted with five military ATCOs in the MicroNav BEST Training Simulator. During each experiment session a complex approach control task, and a surveillance radar approach, were performed and both subjective- and objective data (i.e. simulator and eye-tracking data) was collected. Although the results showed no significant differences between the sessions, new insights into skill decay indicators were gained, including factors such as change in ATCO scan-patterns and the influence of ATCO experience (i.e. level and exposure) on the task.

© 2022 The Authors. Published by ELSEVIER B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the 34th Conference of the European Association for Aviation Psychology

Keywords: Competence Retention; Retention Modelling; Situational Assessment; Eye-Tracking; Skill Decay

1. Introduction

Preventing air traffic controller (ATCO) skill decay is crucial to flight safety. Training is a good remedy, but expensive and sometimes unnecessary as people tend to differ in the extent to which skills are retained over time. Personalisation of training may reduce unnecessary skill loss as well as unnecessary training. In order to personalise retention training it is important to understand how and when skill decay occurs and when it has an impact on safety. How can we predict skill decay and how can this be used to adapt training?

* Corresponding author.

E-mail address: Petra.ten.Hove@nlr.nl (P.E. ten Hove).

Over the last few years several studies were performed on retention of skills and the corresponding skill decay. In a three year research programme, initiated by the Dutch Ministry of Defence, knowledge on Retention and Skill Decay was explored. Vlasblom, Pennings, van der Pal & Oprins (2020) carried out a literature review focussed on competence retention in safety-critical professions. This work served as the basis for the competence retention model reported by van der Pal, Toubman, Crijnen & ten Hove (2021) and van der Pal, Toubman, Crijnen, Binsch, Oprins, de Vries & Landman (2021) depicted in Fig. 1. The aim for the last year of the research programme was to perform preliminary studies on how to apply the model. One of which is the focus of this paper, looking at how to apply this model in the military ATCO domain.

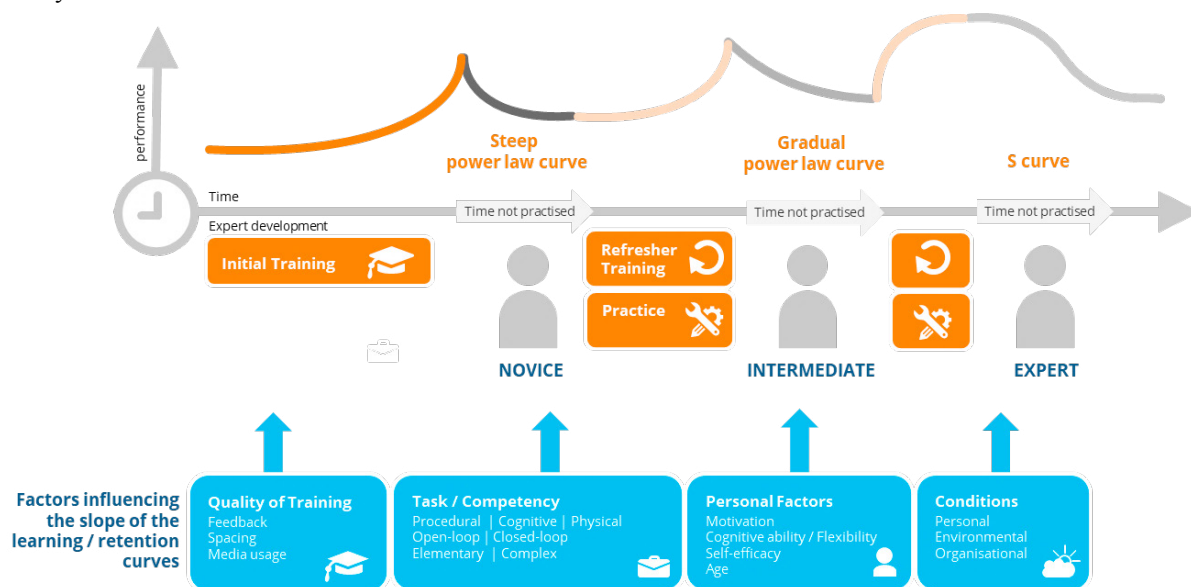


Fig. 1. The competence retention model (NLR)

The competence retention model shows the general predicted performance over time during periods of (non-) practice. The predicted performance is influenced by the experience level (i.e. student, novice, intermediate and expert). In general, the more experience the longer the period of retention and the less influence external and internal factors have on performance. Factors influencing skill decay are divided in four categories: 1) quality of training, 2) task/competency, 3) personal factors and 4) conditions, while Toubman et al. (2022) underlines that predicting retention of skills is specific to the task trained – the second category in the competence retention model. The research presented in this paper used the models and techniques presented by Vlasblom et al. (2020) and van der Pal et al. (2021) in a more practical way. The models were used to assess ATCO competence retention in terms of skill decay and changes in individual ATCO scan patterns.

The goal of this preliminary piece of research was to investigate methods and techniques for determining and predicting ATCO competence retention and skill decay. The underlying reason being personalisation of training and instructor support in predicting when (refresher) training could be necessary. The first step towards predicting skill decay was to determine how performance could be monitored over time for the individual controller. Measurement of performance is domain specific, therefore a standard for ATCO performance assessment had to be determined. The ATC Cognitive Process & Operational Situation (ACoPOS) model (Schuver-van Blanken, Huisman, & Roerdink, 2010) defines ATCO competences in three main cognitive processes: situation assessment, problem solving & decision making and attention management & workload management. To delineate the experiment, the skill ‘situational assessment’ was chosen to examine skill decay in ATCO performance. Both objective- and subjective data were used to determine the effects of this skill decay on the situational assessment of ATCOs. Visual scan behaviour is considered a core element for ATCO situational assessment (Lundberg, 2015), therefore eye-tracking was used to understand

scanning behaviour. Additionally instructor and participant questionnaires were used for subjective data and to determine the effects of skill decay factors in the ATCO context.

2. Method

2.1. Experimental Design

To examine the retention model in a Military Air Traffic Control environment, experiments were carried out. Fig. 2 depicts the experimental design. Firstly, a questionnaire was used to determine the period of non-exposure. This was directly followed by the first experiment session where the performance of (student) controllers was measured in a simulator environment (i.e. a baseline was set). After the first session, controllers received the opportunity to train their skills. This additional training was mandatory for controllers with a below-standard performance (as reported by the instructor). They were required to train until their work was back up to an acceptable level (at least 6 out of 10 on a score sheet). The training period was followed by a period of non-practise of 2-4 months. The second experiment session measured the performance of the controllers again. During the experiments various components of the retention model (Fig. 1) such as performance, time on tasks and influencing factors were examined.

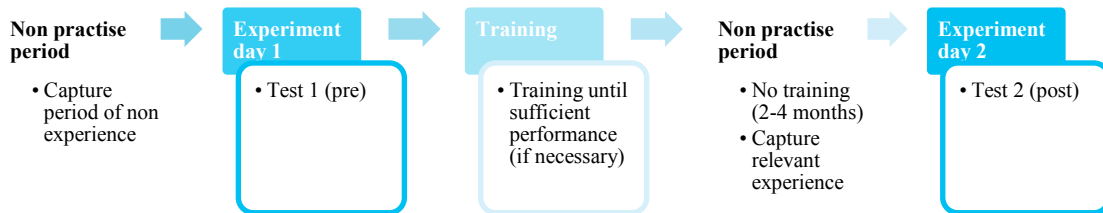


Fig. 2. Experimental Design

2.2. Scenarios

Experiment sessions (i.e. experiment day 1 and 2) consisted of a complex and busy traffic scenario. This scenario used was a version of the most complex examination scenario for this group of military approach controllers. Four different versions (A-D) of the scenario were available, each with a different traffic setup but with equal complexity levels. Each participant randomly received one of the versions to avoid influencing other participants.

At the end of each scenario, an Instrument Landing System (ILS) failure occurred requiring a Surveillance Radar Approach (SRA) procedure. This is a radar instrument approach needing active assistance from the ATCO. In real life an SRA procedure is uncommon according to subject matter experts, and is therefore possibly sensitive to skill decay.

2.3. Participants

In total five (student) military air traffic controllers with an approach control (APS) rating participated in this study. The group consisted of 2 female and 3 male ATCOS and all were aged between 25-65 years old. All participants provided written informed consent to anonymously record and use their data. The group consisted of 1 student controller (waiting for the next stage of training, novice level), 2 recently licensed controllers with operational shifts (intermediate level) and 2 instructors (no operational shifts, 1 novice, 1 expert level).

The five participants were tested on their skills with complex traffic and the SRA procedure. Three of the participants had recent (i.e. less than 3 months ago) experience with busy complex traffic. The remaining two reported that the last time they had dealt with such traffic was 18 months ago. Participants' experience with the SRA procedure differed somewhat. Only two had recent experience (2 and 4 months ago) with this specific approach. The remaining

three reported to have last carried out an SRA procedure 10 months to 2 years ago. Only one participant was exposed to the SRA procedure during the period of non-practise between the experiment sessions.

Table 1. Experience data for all participants

ID	Age	Experience / ATCO type	Before session	Time since complex traffic	Time since SRA procedure	Extra training needed / used
A	30-35	Novice	1	18 months	24 months	No
		Instructor	2	3.5 months	3.5 months	3 sessions
B	60-65	Expert	1	18 months	18 months	Yes, full
		Instructor	2	3.5 months	3.5 months	5 sessions
C	25-30	Novice	1	3 months	10 months	No
		Unlicensed	2	2 months	2 months	None
D	25-30	Intermediate	1	0 months*	4 months	Yes, partial
		Licensed	2	0-3 months*	1 month*	2 sessions
E	30-35	Intermediate	1	0 months*	2 months	No
		Licensed	2	0-3 months*	3 months	2 sessions

*On the job experience instead of simulator session

2.4. Data

To determine the ATCO performance level and if skill decay had occurred, two types of data were used: objective data - from the simulator (i.e. aircraft and label positions on the radar screen) and eye-tracker - and subjective data - from an instructor assessment form and participant questionnaires (see Fig. 3). Due to the small group of participants, no quantitative statistical analysis was performed, only descriptive statistics are reported.

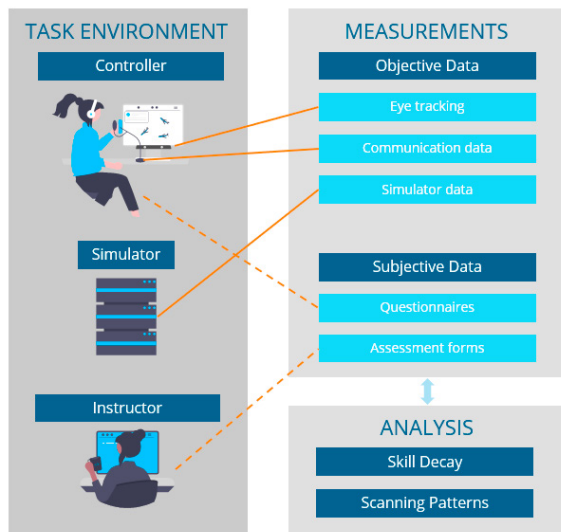


Fig. 3. Data Collection

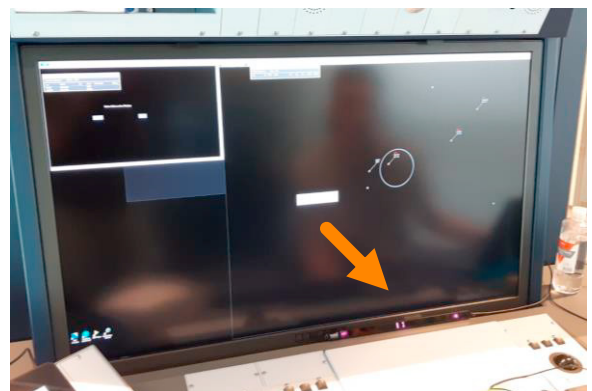


Fig. 4. Eye-tracking equipment installed on BEST radar simulator. Source: NLR

The experiment sessions took place in the radar controller training simulator used for training Royal Netherlands Air Force (RNLAf) Military Air Traffic Controllers: The MicroNav BEST training simulator. The radar environment was set up for Dutch airspace focussing on the Eindhoven (EHEH) military approach control area. (Background) traffic

Index - NASA-TLX (NASA,2006) to track the perceived workload. For further clarification of the questionnaires one-on-one interviews were held with the participants and instructors.

3. Results

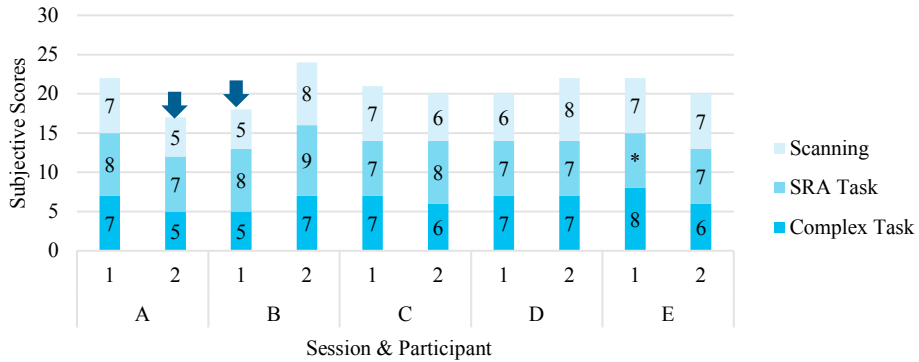
3.1. Subjective Performance Data

To detect skill decay, the first step was to analyse the subjective performance data. Unfortunately few conclusions can be drawn when comparing session 1 and 2. Table 2 shows the average of the task and competency scores given by the instructor for session 1 and 2. On average the participants scored the same in session 1 ($M = 7.0$, $SD = 0.8$) and session 2 ($M = 7.0$, $SD = 0.9$). The raw NASA task load index indicates that the participants perceived a slightly higher task load during session 1 (45/100) than session 2 (40/100).

Table 2. Average subjective scores for the tasks and competencies

<i>Task or Competency</i>	Session 1		Session 2	Mean	SD
Complex Task	6.8	▼	6.2	6.5	1.0
SRA Task	7.4	—	7.6	7.7	0.7
Safety				6.6	1.0
Safety	6.8	▼	6.3		
Skills				6.9	0.8
Analysing	6.6	▲	7.0		
Attention Management	7.0	—	7.0		
Planning	7.0	▼	6.6		
Prioritization	7.4	▼	6.6		
Scanning	6.4	▲	6.8		
Work & Mental Pace	7.2	▼	6.8		
Cognition				7.1	0.9
Knowledge & Understanding	6.8	▲	7.4		
Communication				7.2	0.6
Coordination	7.2	—	7.2		
Radio Telephony	7.2	—	7.2		
Speaking & Listening	7.2	—	7.2		
Attitude				7.2	0.6
Collaboration	7.0	—	7.3		
Decision Making	7.4	▼	6.8		
Responsibility	7.4	—	7.4		

Individual participants' subjective performance data does however reveal some differences. Fig. 5 depicts the three main subjective performance scores per participant as assessed by the instructor for 1) the complex task, 2) the SRA task and 3) the competency scanning. Skill decay or skill growth seems to be dependent on the background and current activities of the participant. Two sessions (of candidate A and B) showed scores below standard (below 6 out of 10). Assuming the candidates scored sufficiently in the past on these type of exercises to pass their former training, session A2 and B1 were marked as 'skill decay' and analysed in more detail.



*Due to simulator issues session E1 was aborted, so no SRA procedure was assessed.

Fig. 5. The main subjective assessment scores per participant

Interestingly, a difference can be noticed between the experience levels of the controllers. The performance results in Fig. 6 seem to match the curves of the ‘retention model’. Indicating the participants with a novice level are sensitive to skill decay after a short period of time (steep power curve). While the participant with an expert level of experience showed skill decay after 2 years of no practice but is quickly back on track (S curve).

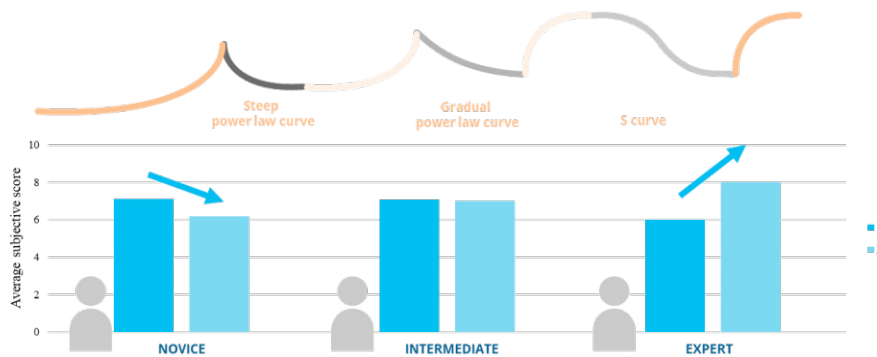


Fig. 6. Average of main subjective assessment scores of session 1 and 2 between the experience groups

Subjective indications of skill decay became clear from the interviews and questionnaires. Some participants indicated a feeling of skill decay, for example, they said “I feel a bit rusty”. Others commented that they had focussed less on efficiency and “played it safe” to prevent performance reduction. When addressing moments of performance decay during interviews, several of the instructors and participants reported an unstructured scan pattern. They let go of their scan cycle and/or focussed on the conflicting flights while missing others. Additionally, some reported that their planning and coordination fell behind, which was also observed in their communication, i.e. Radio Telephony (RT) calls became shorter or small mistakes were made such as using the wrong callsign.

3.2. Objective Performance Data on Situational Assessment

Investigation of the objective data can provide a better understanding of skill decay indications in situational assessment. Table 3 presents the main results from the objective performance data. The perception score (based on the algorithm described in section 2.5) for participant A shows a decrease, and for participant B an increase between the sessions, in line with the observed skill decay. However, missing data and constraints from the eye-tracker (described later on in the limitations section) make it difficult to draw clear conclusions, especially for the other participants. The last column shows the availability of eye-tracker data, which provides an indication of the quality of the data.

Table 3. Objective simulator & eye-tracking data combined

ID	Session	Perception*	Mean Time Anticipation*	Missed Flights	Total Flights	Eye-tracking available
A	1	43	0:02:01	14	50	90%
	2	35	0:02:24	10	48	65%
B	1	37	0:03:16	11	49	90%
	2	42	0:04:09	11	47	87%
C	1	29	0:03:06	14	50	55%
	2	32	0:03:22	15	54	67%
D	1 **	48	0:01:00	15	30	89%
	2	33	0:03:51	10	47	93%
E	1 **	-	-	-	-	-
	2 **	58	0:03:03	1	17	85%

* Perception and anticipation score algorithms as described in the method section

** Data incomplete

After analysing the objective data, two noticeable things were found. Firstly, one session (A2) took considerably more instructions (31 calls versus 16 on average) to complete the SRA procedure (see Fig. 7). That being said, the subjective SRA score was not below standard, but it is a sign of more effort resulting in an overall lower performance score.

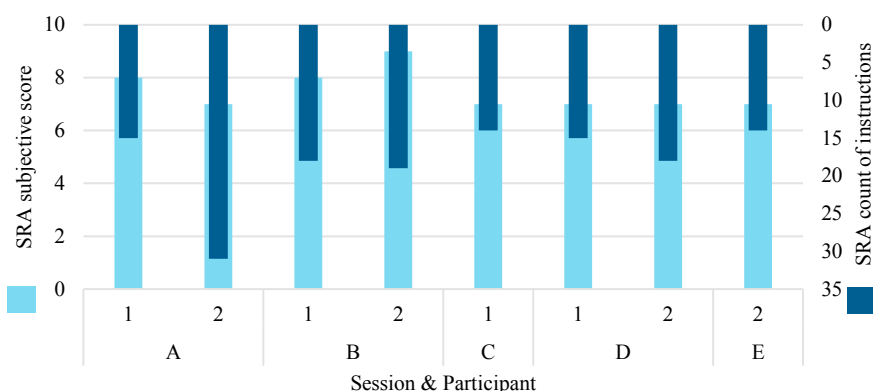


Fig. 7. Performance during SRA procedure

Secondly, after analysing eye-tracking results from participant A and B (that were marked with skill decay) some differences in scan pattern were noted. Participant A tends to show a smaller scan pattern in session 2 - where skill decay occurred. As can be seen in the broad (Fig. 8a) and small (Fig. 8b) scan pattern inside and outside the area of responsibility. This could be an example of the tunnelling effect, where an ATCO only focusses on conflicts without completing the scan cycle. While it is actually important for an ATCO to continuously scan all traffic to maintain situational awareness (Lundberg et al, 2014). On the other hand, participant B’s scan pattern in session 1 - where skill decay occurred - seems less clear. The controller themselves described the first session as more ‘chaotic’ (e.g. Fig. 9a) in comparison to the second, improved session (Fig. 9b). The controller explained this as ‘losing his scan cycle’, however such a sequence is difficult to show by a heat map (Westin et al., 2019). These findings relate to the lower perception scores for participant A and B, however as noted earlier this cannot be generalised to the other candidates.

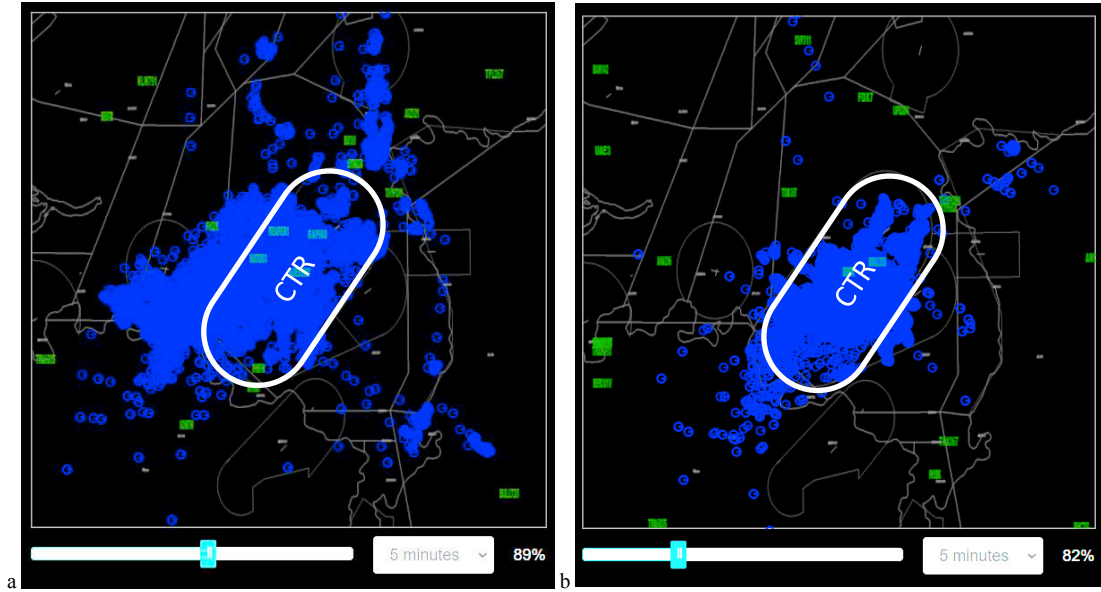


Fig. 8. Eye gaze during 5 minutes in the middle of the scenario for session (a) A1 & (b) A2

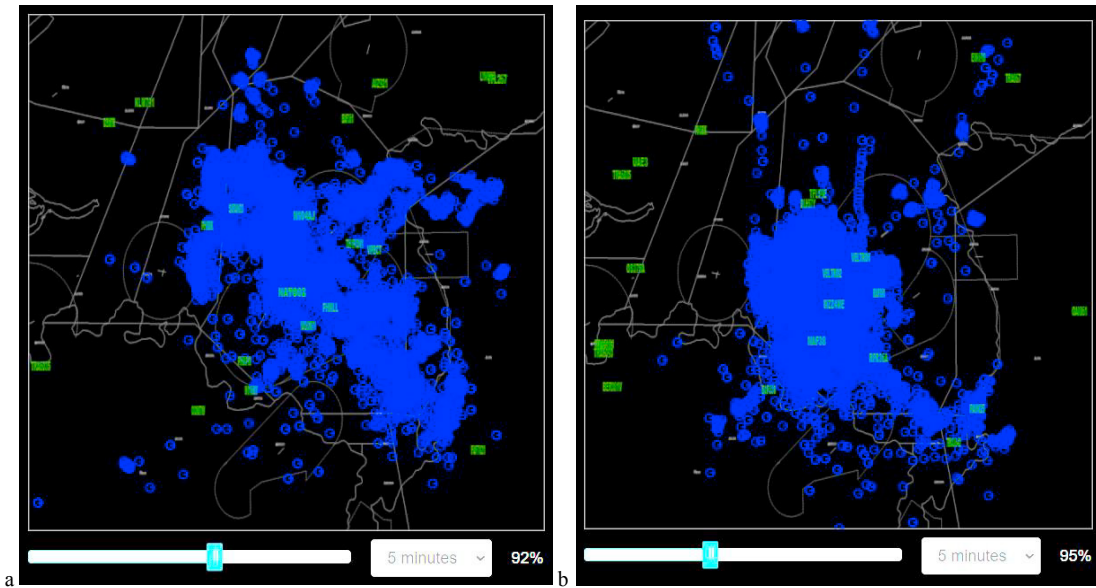


Fig. 9. Eye gaze during 5 minutes in the middle of the scenario for session (a) B1 & (b) B2

3.3. Skill Decay Factors

In order to understand how the competence retention model (Fig. 1) can be applied for ATCO training we delve into the skill decay factors: 1) quality of training, 2) task/competency, 3) personal factors and 4) conditions.

No comparison could be made on the ‘quality of training’ factor. All participants (5/5) rated both their initial training and the feedback received as ‘high quality’(on a scale of 6). Only one person received retention training before the experiments and also rated it as high quality.

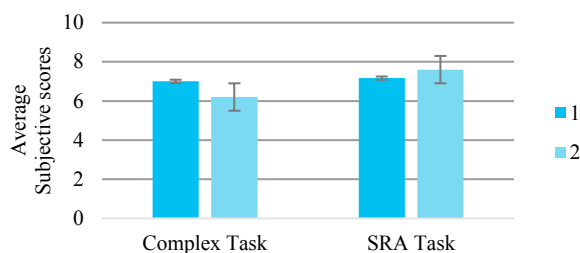


Fig. 10. Average main subjective scores of complex traffic and SRA handling between the sessions

According to the *'task/competency'* factor, a difference in the subjective performance scores is noted between the 'complex traffic' (complex/cognitive) task and the 'SRA' (procedural) task (see Fig. 10). The complex task shows slight decay while the SRA task shows slight growth in the subjective scores. This could implicate that complex skills are more difficult to keep on track compared to procedural tasks. Looking in detail at the two sessions with skill decay, the competency score for 'scanning' and 'safety' showed the most decay. Possibly these competencies are more sensitive to skill decay or skill decay indications of ATCOs are the first to be visible in these competencies

The intermediate experienced controllers reported in both sessions difficulties with the simulator system and environment, because of dissimilarity to their daily operations. This confirms that procedural skills are sensitive to skill decay (Vlasblom et al., 2020).

For *'personal factors'* no major findings can be reported. Four participants scored their motivation as high, one as neutral. Participant age varies between 20-30, except for one participant who was over 60. Although there is too little data to analyse the age factor, it could be considered of interest that the older participant mentioned *"it costs me more effort to keep my level of performance up, than it would have some years ago"*.

Within the *'conditions'* category an effect was noticed for *'personal conditions'*. In 3 of the 10 sessions the participants got less sleep the night before and two of them reported they were less alert during the experiment. This only seemed to influence the novice participant (A2) who performed that session below standard (subjective assessment). This underlines that novices may be more sensitive to skill decay factors such as (personal) conditions (Vlasblom, Pennings, van der Pal & Oprins, 2020).

4. Conclusion

The research question in this paper is twofold. The first step was to determine how performance deterioration, for individual controllers, could be monitored over time. The second step was to determine the effects of skill decay factors in the ATCO context.

With regard to the first step, indications for monitoring skill decay over time were found. For example, small mistakes in procedures, a shorter length of communication or an increase in instructions. Analysing the scan pattern could be a suitable indicator for skill decay in early stages, in particular for situational assessment. Note that there are several ways that skill decay can manifest itself, such as smaller scan pattern, more chaotic, or missing scan cycles. Multiple measurements are necessary for the detection and prediction of skill decay.

With regard to the second step, the main factor influencing skill decay is the experience level of the controller. According to the experiment, no skill decay is expected during an interval of at least four months for controllers with an intermediate or expert experience level. For novices, on the other hand, skill decay could occur quickly. This is why an ATCOs exposure to tasks, experience level, and amount of training are all important factors for skill decay prediction models. These results are in line with the literature and theoretical model on competence retention (Fig. 1).

In the experiments some indications for other factors were also observed. The performance deterioration was different per type of task and person. Personal conditions such as amount of sleep and stress seemed to negatively influence the performance of novice controllers.

At the time of writing, data and retention modelling is not ready-to-use, however, it is still possible to implement retention strategies from the competence retention model in initial and continuation training. This study supports the following guidelines for ATCO (retention) training:

- Keep in mind that environmental and personal factors have more influence on performance of students and novices.
- Repeat and distribute training to make for better skill retention
- Provide more support, training and feedback for novice controllers.
- Offer more training possibilities for elderly controllers when changes occur in the task, especially procedural changes.

4.1. Limitations

It should be noted that this research had many methodological challenges including the limited number of participants, which means that the findings cannot be backed by quantitative statistical analysis. Therefore it is important to continue skill decay research for ATCOs and to keep in mind that the results of this research may not be generalised.

In addition, experienced operational controllers were unavailable to participate in the experiment. Backed by theory, we expect experienced controllers to show no or little skill decay, because of their daily operation in a complex air traffic environment combined with their level of experience. This complicates the detection and research of the skill decay concept in an experiment session with experienced operational controllers.

Eye tracking data has shown to be a valuable measure to provide instructors with insight and understanding of scan patterns and controller tactics. With this insight the instructor is able to provide more targeted feedback, which benefits the training process. However, before eye tracking can become widely available, certain obstacles must be overcome.

- 1) The eye tracker was positioned below the radar screen (which is the suggested location). This radar screen is a large and wide monitor. In this particular setup the eye tracker seemed to be not always accurately track the controllers eye gaze. Especially near the corners and upper part of the screen. This inaccuracy is currently influencing the learning algorithms, e.g. this influences the reaction time in detecting flights near the edges.
- 2) The eye tracker requires direct sight of the eyes, which can sometimes be obstructed by hands or objects (e.g. holding a pen).
- 3) During the experiment the availability/quality of the eye tracking was logged. Analysing the results it became apparent that the quality of the eye tracking was strongly dependent on the amount of movements that a controller made. For example, instances were reported that the eye gaze was lost when the controller moved forwards and backwards in their seat.

Taking the above into consideration it is important to understand that this is a preliminary study meant as an initial exploration of retention and skill decay in ATCOs. Substantially more research is needed in the field of skill decay and retention modelling before it can be put to use in daily operations.

Acknowledgements

We would like to thank the Dutch Ministry of Defence for funding this research within the program Education and Individual Training (Grant number L1806). We are also very grateful to the School of Air Control (Royal Netherlands Air Force) for enabling the experiment and MicroNav for their support on extracting the simulator data.

References

- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.
- Hove, P.E. ten, Mohrmann, J.F.W., Bos, T.J.J., Apeldoorn, M. van, & Muynck, R. de (2020). *Rapportage KAV ANSP's 2020 Learning Analytics* (Report No. NLR-MEMORANDUM AOTS-2020-030). Internal Report Netherlands Aerospace Centre.
- Lundberg, J. (2015). Situation awareness systems, states and processes: a holistic framework. *Theoretical Issues in Ergonomics Science*, 16(5), 447-473.
- Lundberg, J., Johansson, J., Forsell, C., & Josefsson, B. (2014, July). The use of conflict detection tools in air traffic management: an unobtrusive eye tracking field experiment during controller competence assurance. In *Proceedings of the International Conference on Human-Computer Interaction in Aerospace* (pp. 1-8).
- Mohrmann, J.F.W., Hove, P.E. ten, Bos, T.J.J., Apeldoorn, M. van, Schaik, J. van, Jansen, R.R.J., & Tillema, G.H.J. (2021). *ANSP Learning Analytics, Rapport 2021* (Report No. NLR-TR-2021-431). Internal Report Netherlands Aerospace Centre.
- NASA. (2006). NASA TLX: task load index. Retrieved from: <http://humansystems.arc.nasa.gov/groups/TLX/>
- Pal, J. van der, Toubman, A., Crijnen, J.A., & Hove, P.E. ten (2021). Exploring competence retention and retention modelling (Report No. NLR-TP-2021-257). Internal Report Netherlands Aerospace Centre.
- Pal, J. van der, Toubman, A., Crijnen, J.A., Binsch, O., Oprins, E., Vries, D. de, & Landman, A. (2021). Retentie. In: E. Oprins, D. de Vries, A.-F. Holster, M. 't Hart (Eds.) *Flexibel leren en Ontwikkelen bij Defensie en de politie* (pp. 133-156). Internal book TNO Soesterberg
- Schuver-van Blanken, M., Huisman, H., & Roerdink, M. (2010). The ATC Cognitive Process & Operational Situation Model. In European Association for Aviation Psychology Conference. Budapest, Hungary.
- Toubman, A., Oldenbeek, M. van & Claessen, O. (2022). Predictive modelling of retention. (Report No. NLR-TR-2021-130). Internal Report Netherlands Aerospace Centre.
- Vlasblom, J.I.D., Pal, J. van der, Suijkerbuijk, H.C.H., & Leij, O. van der (2020). *Learning Analytics for the School of Air Control* (Report No. NLR-CR-2020-147). Internal Report Netherlands Aerospace Centre.
- Vlasblom, J.I.D., Pennings, H. J., Pal, J. van der, & Oprins, E. A. (2020). Competence retention in safety-critical professions: a systematic literature review. *Educational Research Review*, 30, 100330.
- Westin, C., Vrotsou, K., Nordman, A., Lundberg, J., & Meyer, L. (2019). Visual scan patterns in tower control: Foundations for an instructor support tool. In *SESAR Innovation Days*.