Analysis of learning curves in on-the-job training of air traffic controllers

Esther Oprins
TNO Defense, safety & security, Soesterberg, The Netherlands
esther.oprins@tno.nl

Ernst Burggraaff
Air Traffic Control the Netherlands (LVNL), Amsterdam, The Netherlands
e.r.burggraaff@lvnl.nl

Robert A. Roe
Maastricht University, Dep. of Organization & Strategy, Maastricht, The Netherlands
r.roe@maastrichtuniversity.nl

Introduction

This chapter describes a competence-based assessment system, called CBAS, for air traffic control (ATC) simulator and on-the-job training (OJT), developed at Air Traffic Control The Netherlands (LVNL). In contrast with simulator training, learning processes in OJT are difficult to assess, because the learning tasks cannot be planned in advance due to the ongoing air traffic. The assessment system in OJT was designed in such a way that the trainees' progression can nonetheless be monitored. The reliability and validity of CBAS have been evaluated in previous research (Oprins, Burggraaff & Van Weerdenburg, 2006, 2008; Oprins, 2008). Here we present an evaluation with regard to the analysis of learning curves derived from assessment results. An adequate assessment of learning processes showing differences in individual learning patterns (e.g., slow starters, learning plateaus) and in performance (strengths and weaknesses) is of crucial importance as a basis for feedback on the learning process and for pass-fail decisions. CBAS compares the trainees' actual performance to required performance at successive moments of time. Under the assumption that performance is a reflection of the learning process (Oprins, 2008), it extracts learning curves from a sequence of performance measures over time. If trainees are learning, then their performance will increase (Oprins, 2008).

In this chapter we start with describing the main principles of CBAS and the use of learning curves in OJT. Next, we explain the method by which learning curves are derived from the assessment results. Finally, we present the results of the analysis of learning curves.

The competence-based assessment system (CBAS)

The notion of competence refers to the successful integration of knowledge, skills and attitudes and their application in realistic environments (Oprins, 2008). Competences indicate an individual's ability to effectively perform certain tasks when situational factors are held constant. As competences are based on learning, the assessment of competences at a particular moment in time is essential for adequate feedback, needed to improve individual performance. The first step in the design of the competence-based assessment system (CBAS) at LVNL was a competence analysis based on input from a peer group of air traffic controllers and literature research. The way in which the competence analysis was carried out has been described elsewhere (Oprins, Burggraaff & Van Weerdenburg, 2006). Here we describe the resulting ATC Performance Model which has served as the framework leading the design of CBAS (see Figure 1).

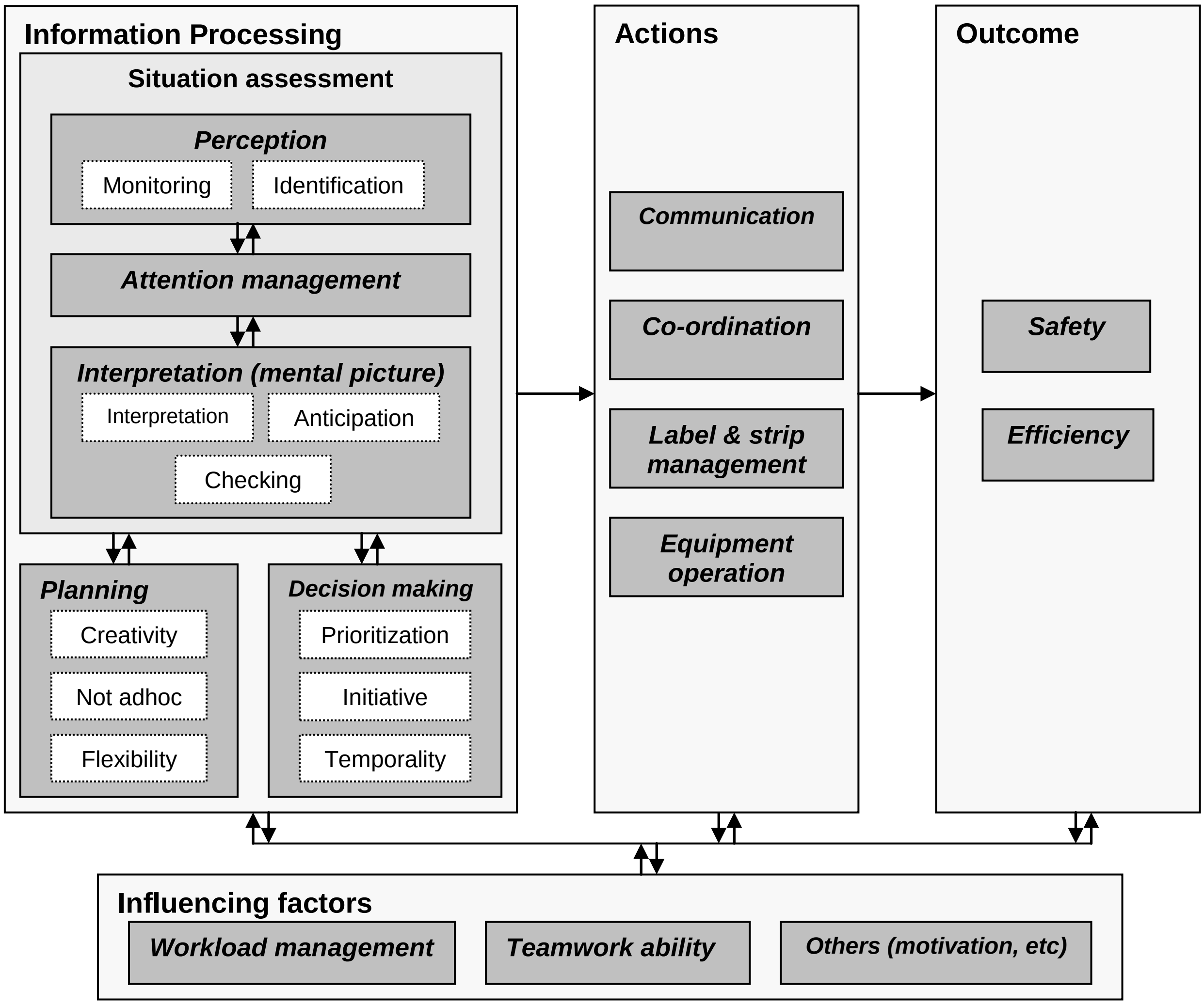

Figure 1. ATC Performance Model

In CBAS, each competence is represented by a set of *performance criteria*. The criteria are formulated in the jargon of the air traffic controllers in order to maximize comprehension and recognition of behaviours. They have been given the format of 'behavioural markers' (O'Connor, Hormann, Flin, Lodge & Goeters, 2002) and are rated at a six point rating scale. The use of these performance criteria is important as they allow controllers, in their role of assessors, to express their 'gut feeling', and to specify why a trainee performs (in)sufficiently and what should be done by the way of individual training interventions.

During training the same competences are assessed a number of times, in order to follow trainees' progression over a certain time period and to gain a better insight in deficiencies in performance in various task situations. In ATC training competences are assessed against augmenting *performance standards.* More specifically, simulator training and OJT are divided into phases, each having its own standards. The standards have been formulated as 'behavioural examples', which resemble the 'behaviour anchors' used in behaviourally anchored scales (BARS; see Berk, 1986). However, they do not specify scale positions but standards to be achieved at the end of each phase. In this way, it is clearer to assessors what is expected from trainees in intermediate phases. For the trainees it is clearer which competences they have to develop further in a specific phase. While the phases in simulator training have mainly been defined by the sequence of simulator exercises, there is no fixed structure in OJT due to the ongoing live traffic. The OJT is divided into four phases, following three main principles: (a) the requirements during training move from *safety* towards *efficiency*, with (b) increasing *traffic complexity*, and with (c) decreasing *aid of the coach* because of acquired expertise. The phases are designated as:

1. Familiarization phase
2. Learning phase 1
3. Learning phase 2
4. Consolidation phase.

The length of each phase is flexible, dependent on trainees' progression. See Table 1 for an example of *building a mental picture* in OJT of area control (ACC). Five *performance criteria* (see left column) are rated at a six points scale for assessing the competence *building a mental picture*. The *performance standards* that belong to each phase, i.e. familiarization phase, learning phase 1 and 2, and the consolidation phase, are presented as behavioural examples in the next four columns. The aformentioned three principles can be recognized in the examples.

To get a complete picture of trainees' performance and to measure progression over time, CBAS uses continuous assessment of performance in retrospective progression reports, using the performance criteria of the ATC performance model. The progression reports are filled in after one to two weeks of training. Multiple assessors are involved in both types of assessment for maximizing reliability. A web-based assessment tool is used to fill in progression reports, to store the results in a database, and to generate overviews of trainee performance – information that is useful for monitoring trainees' progression. Pass-fail decisions are based on the trainees' progression over time. Trainees go on to the next phase if all performance ratings are rated with a value 4 or more which serves as the required standard. Trainees can stay longer in a phase if necessary, but they fail if they do not show any progression over a longer period of time. Pass-fail decisions are not based solely on a quantitative measure yet, but they are based on expert judgement of training managers. Continuous assessment with rating the same competences at subsequent moments in time makes it possible to derive learning curves from the assessment results.

Table 1. *Performance standards for* building a mental picture *in ACC OJT.*

Building a mental picture	Familiarisation Phase	Learning Phase 1	Learning Phase 2	Consolidation Phase
Keeps a clear overview of the traffic situation by scanning regularly **Looks, observes and takes action if necessary** **Checks available information to be correct** **Guards the identification process of the label presentation** **Anticipates future and variable traffic situations**	As the trainee is still getting used to live surroundings he is inclined to focus too much on certain conflicts or flights instead of scanning the entire sector regularly for other traffic requesting attention. Therefore it may happen that certain flight is not cleared further. He may also be surprised by the call of an aircraft. At this stage the trainee finds it hard to react at everything he sees and hears. There is more visual and audible information than in the simulator. Therefore it is difficult to act adequately when traffic gets heavier. The trainee finds it hard to estimate intermediate FL's, the sequence of aircraft (e.g. diverging/converging tracks, speeds, and influence of the wind). At this stage the trainee finds it hard to divide his attention well and have a complete image of the traffic at all times, as he still has to get used to the live surroundings. Therefore it may happen he doesn't discover a flight without a label in time. And when he fits the strip into the sequence he doesn't realise the SSR code has to be adapted. Sometimes the coach has to remind him to change the SSR code or have it changed. ===================================== The coach is closely involved in the handling of traffic and gives tips regularly for a safe and efficient traffic flow. The coach asks many questions like" what will you do with that information, what does this information mean", etc.During heavier traffic the coach can take over control completely for a while.	During normal traffic the trainee has a good picture of the traffic situation as he scans regularly. He is no longer surprised by calls of any aircraft as he doesn't only focus on specific conflicts or flights, but scans the entire sector regularly. Only during complex traffic may he still be inclined to loose his overview. By watching changes in information closely, listening to calls by aircraft processing information from the sector(s) and taking action where necessary the trainee is in control during normal traffic situations. He is able to concentrate during a longer period of time and is pro-active. The trainee is monitoring all the time checking that the traffic situation develops as expected, e.g. after giving certain instructions, checking the wind influence and the pilot's reaction to the instructions. Only in complex situations the trainee finds it hard to estimate intermediate FL's, the sequence of aircraft (e.g. diverging/converging tracks, speeds, influence of the wind). During normal traffic situations the trainee can create a good mental picture of all traffic, also of flights that have not been identified by the system. He links the flightplan to the radarposition in time and correctly, the more so as he has remarked on the strips that the SSR code has to be adapted. ===================================== The coach still gives tips to increase efficiency and only has to intervene in complex situations when safety is at stake.	During normal and complex situations the trainee has a good overview of the traffic because he is scanning regularly. By watching changes in information closely, listening to calls by aircraft processing information from the sector(s) and taking action where necessary the trainee is also in control during complex traffic situations. He is able to concentrate during a longer period of time and is pro-active. The trainee has no problems estimating intermediate FL's, the sequence of aircraft (e.g. diverging/converging tracks, speeds, and influence of the wind). During any traffic situation (also complex) the trainee can create a good mental picture of all traffic. Also of flights that have not been identified by the system. He links the flightplan to the radar position in time and correctly and sees to it that every flight has the correct SSR code. ===================================== The coach gives tips only occasionally for efficiency reasons and intervenes only rarely during complex situations and if safety requires.	The trainee works independently as stated in Learning phase 2 and has acquired sufficient experience at the end of this phase to take the practical exam. ==================== The coach is only present as the person who is formally responsible for the safety.

Learning curves in CBAS

Learning curves and learning theory

Learning curves are usually presented as growth curves measuring performance of the same task execution at successive moments in time. Their purpose is modelling learning processes. A simplified ATC task, the Kanfer-Ackerman task, has often been used for examining complex skill acquisition (e.g., Ackerman, 1989; Lee & Anderson, 2001; Taatgen & Lee, 2003). General learning theory says that each learning curve ends in an asymptote (cf. learning plateau) in conformance with the power law of practice (Newell & Rosenbloom, 1981), see Figure 2.

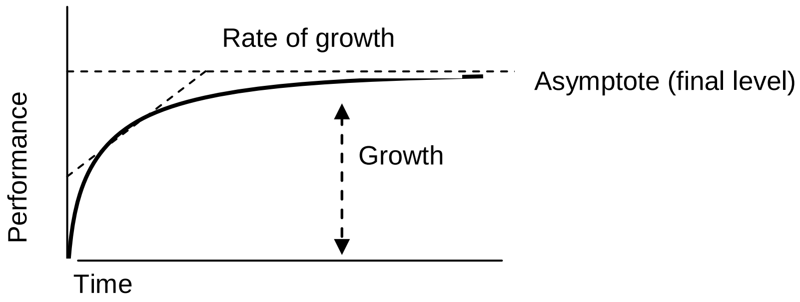

Figure 2. Growth curve of learning

The growth curve of learning as shown in Figure 2 especially applies to simple skill acquisition. Learning curves can take different forms depending on task complexity, task consistency, and individual differences in learning. In the case of complex tasks learning curves are different. Complex tasks pose higher demands on cognitive abilities than simple tasks (Ackerman, 1989) and they typically comprise consistent and non-consistent task components (Ackerman, 1989; Schneider, 1990). Consistent task components show large improvements through practice whereas non-consistent tasks do not (Schneider, 1990): consistent tasks require automation while non-consistent tasks require more controlled information processing. A trainee needs time to assimilate new knowledge and skills with previous experiences in order to automate skills. As a result, the learning process of each individual has asymptotes or intermediate learning plateaus. The ATC task can be subdivided into many task components. Under the assumption that these task components obey the power law of practice, and that they are learned one after each other, we would expect that the overall learning curve would consist of a sequence of smaller learning curves for each task component to be learned in accordance with the findings of Lee and Anderson (2001). This is visualized as the learning curve in Figure 3.

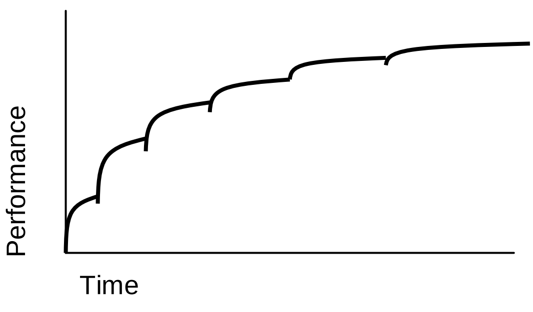

Figure 3. Learning curve of complex skills

However, in on-the-job training skill acquisition is likely to be more complex. Learning is strongly dependent on the quality of coaches and other influences such as the mental pressure to succeed training or the complex working environment (physical conditions, colleagues, etc.). In OJT learning tasks are usually not delivered in a pre-structured sequence. Also, many tasks are trained simultaneously as trainees should be able to manage multiple tasks at the same time. Learning curves would therefore be more smooth than visualized in Figure 3, because they can be seen as the result of many overlapping curves. In addition, the order and tempo of learning may differ across trainees because of individual differences in underlying factors (cognitive abilities, learning styles, personality, pre-education, external influences etc.). Figure 4 illustrates how some variations of learning curves, more smoothed than presented in Figure 3, are expected in complex skill acquisition in OJT with live traffic, under the assumption that all trainees start a the same (zero) performance level.

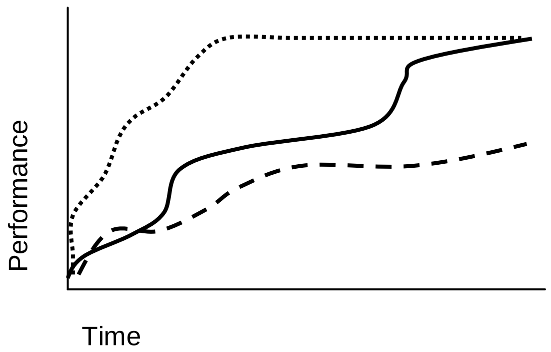

Figure 4. Three variations of learning curves

Figure 4 presents a fast learner who achieves the final performance level earlier than on average (dotted, highest curve), a so-called 'slow starter' who needs more time and who shows an intermediate learning plateau (continuous, middle curve), and a learner who never achieves the required level and probably fails (striped, lowest curve). Ideally, the assessment system should optimally reflect the kind of learning curves depicted here, but our assessment system cannot produce these learning curves in the same way.

Recalibrated learning curves

The learning curves produced by CBAS are based on a weighed sum of competence ratings, however, they are not growth curves as commonly applied. First, the standards against which the trainee is assessed are constantly changing, while remaining mapped on the same six point rating scale. This means that this rating scale, with a value of 4 or more being sufficient, is constantly being recalibrated. When ratings would stay 'sufficient' over time, this implies that the trainee succeeds to meet the increasing standards and shows progression. Second, the recorded measurements are not really continuous since they only reflect performance at specific moments. In between these moments distinct learning processes can take place which cannot be captured completely. In addition, the intervals of measurement differ across trainees due to different training schedules. In this study, the moments in time are presented as rank orders to be able to compare learning curves across trainees. For these reasons, the 'learning curve' produced by our assessment system can only be seen as a derivative of the real learning curve. Figure 5 visualizes these recalibrated learning curves as produced by CBAS.

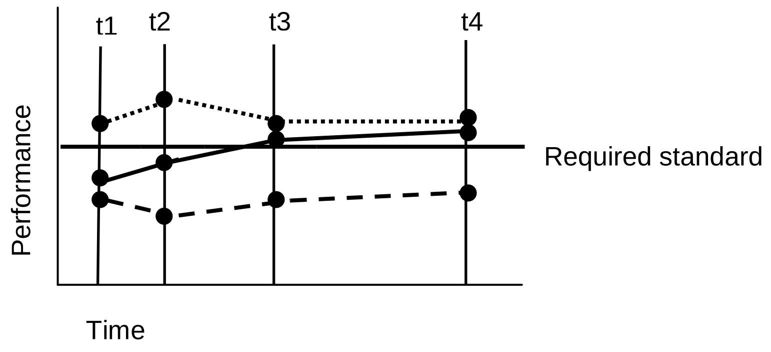

Figure 5. Recalibrated learning curves, produced by CBAS

The dotted, highest learning curve represents a trainee who constantly performs above standards, the continuous middle one reflects a trainee with a temporary learning plateau around t1 and t2 ('slow starter'), and the striped lowest one refers to a trainee who constantly performs below standards and fails. Straight lines connect the points of measurement to obtain a certain learning curve, but these lines do not necessarily represent the real learning processes in-between. In Figure 6 the three trainees are assessed at the same four moments to illustrate possible variations, but in reality trainees are not assessed simultaneously and the intervals differ.

Analysis of learning curves

Goal

The analysis of learning curves, as part of the evaluation of the assessment system CBAS, focuses on the representativeness of the learning curves for learning processes. If the assessment results represent learning processes optimally, then CBAS could be applied as an instrument to gain insight in learning processes of individuals. Adequate feedback could be given, useful interventions could be made (e.g., optimal task selection, coaching and remedial teaching), and pass-fail decisions could be more valid. Therefore, the main goal of the analysis is to find patterns in the learning curves which are representative for learning.

The assessment system should not only be able to represent general learning processes but also competence development over time. Singular competences are even more important for supporting individual learning: they help to identify deficiencies of trainees required for appropriate feedback and interventions. Trainees may differ in the development of specific competences over time. Some competences may be more trainable than others. Thus, we did not only explore patterns in general performance, but also in the various rated competences.

Design

In order to test whether the recalibrated learning curves differentiate between trainees with varying degrees of training success, we designed a study in which expected ‘prototypical’ learning curves were compared with actual learning curves. Three training managers of LVNL made an a priori classification of trainees in *high performers* (passed without problems), *moderate performers* (passed with difficulties), and *low performers* (failed), based on expert judgment. For each class of learning success we defined prototypical learning curves which serve as hypotheses in the analyses. Figure 6 presents these prototypical learning curves in terms of expected zones for the three types of trainees.

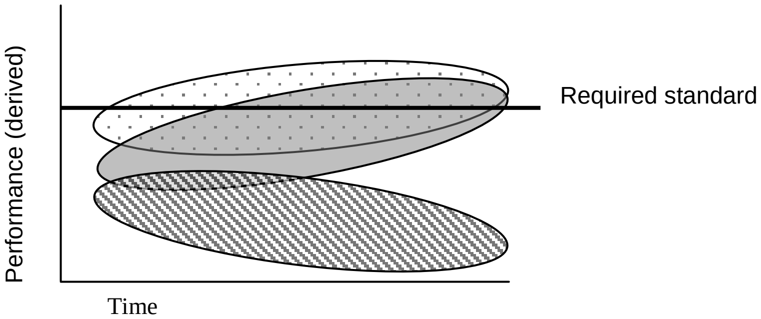

Figure 6. Expected zones for learning curves of high (dotted), moderate (grey) and low (striped) performers

The entrance level of trainees varies because of individual differences. Trainees in the low performance category are likely to start at a lower performance level because they did also performed lower in the preceding simulator training. The high and moderate performers, who both pass training, are expected to achieve the final standards in the end in contrast with the low performers or failures. This latter group is expected to show decreasing performance compared to the increasing standards, the cumulative nature of the learning process, and a lack of self-confidence as a result of their low performance. Because trainees are assessed against phase level and transfer to the next phase when all competences are rated as sufficiently, this picture is expected in each learning phase of OJT.

Method

We used 403 progression reports made for 27 trainees in OJT of area control (ACC). We did the same analyses only for learning phase 1 and 2 because in these longest phases trainees learn most. We visually examined the learning curves of trainees, classified into the three groups, to investigate patterns in learning processes. We did a discriminant analysis to check whether the classification into the three groups was correctly predicted in a quantitative way. Therefore, we defined four variables, divided into two groups, because we did not know yet which measure would be most suitable for defining learning curves quantitatively:

- <u>Performance level</u>: 1) *mean performance level* (weighted sum of competence ratings); and 2) occurrences of *insufficient performance* (value of this weighted sum < 4)
- <u>Progression:</u> 1) *growth* (final minus initial performance level); and 2) *rate of growth* (beta coefficient of the linear regression model, obtained by curve fitting).

An analysis of variance (ANOVA) between the three groups, separately for learning phase 1 and 2, was mentioned to provide insight into the differences in means of the variables across the groups.

In addition, we examined to which extent the three groups differ in means on each competence, both phases together, visually and with an analysis of variance (ANOVA). We also calculated rank order correlations between the variable ‘time’ and the competence ratings for each group to investigate the differences in progression, separately for two phases in OJT.

Results

<u>Learning curves</u>

The derived learning curves for the three groups resemble the prototypical learning curves as presented in Figure 6 to a certain extent. See Figures 7 and 8 for respectively learning phase 1 en 2.

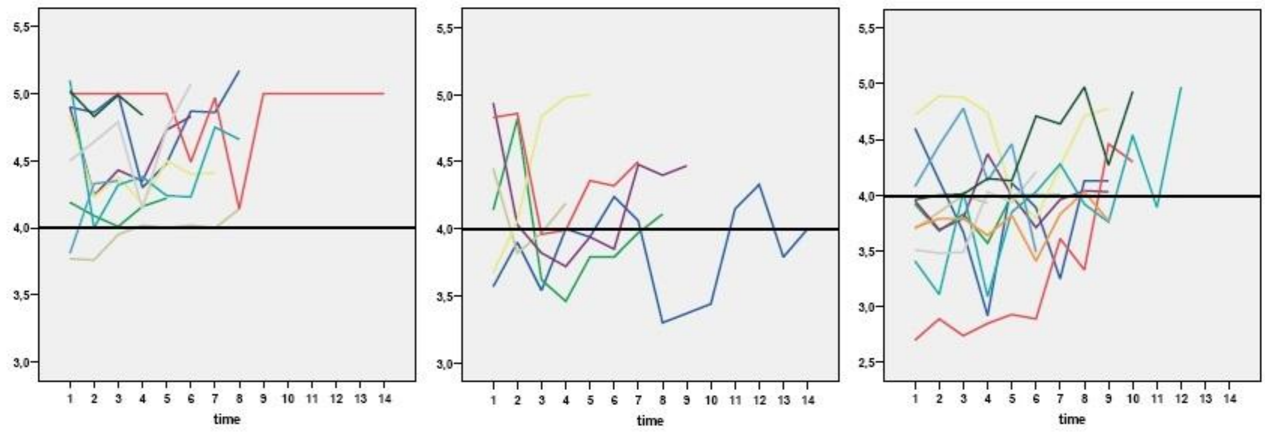

Figure 7. Derived learning curves for respectively high, moderate and low performers in learning phase 1

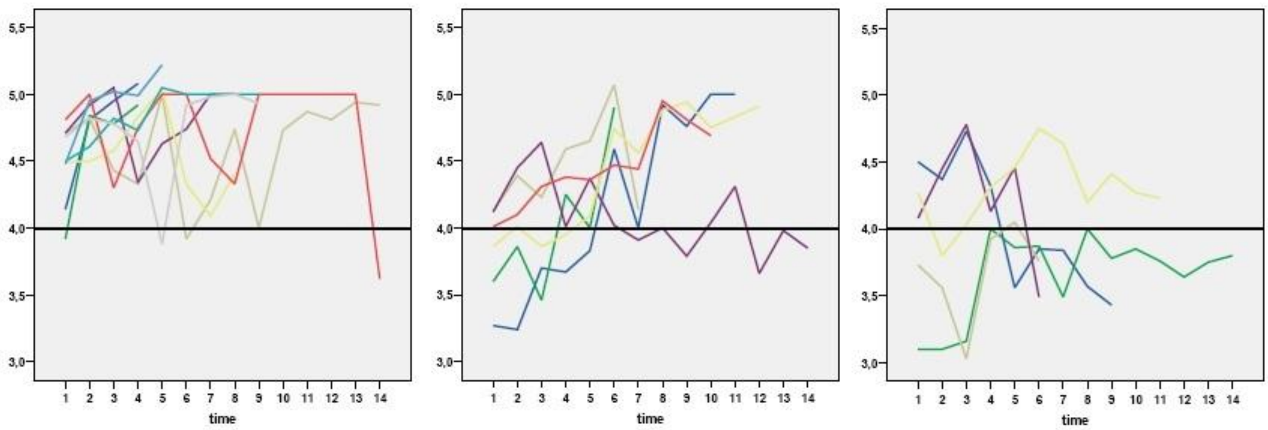

Figure 8. Derived learning curves for respectively high, moderate and low performers in learning phase 2

The graphs for the three groups are quite similar for learning phase 1 and 2 as expected, because the same assessment and training methods are used. The group of high performers (N=10) performs rather constantly above the sufficient standards. We would expect an increase towards the required standards, but this is only visible for moderate and low performers. Apparently, some competence ratings can be insufficient in the beginning, but the weighted sum of ratings stays sufficient for high performers. As expected, the variation between trainees within the group of moderate performers (N=6) is really high. Trainees perform around the standards but with many peaks and downs. It should be noticed that the low performers (N=11) are the same trainees in learning phase 1 and 2; some trainees fail in phase 1 already but others transfer to phase 2 and fail finally. This explains why the results are more positive for phase 1 and why a smaller number of learning curves is presented for phase 2. However, all graphs present some outliers who do not fit in the patterns well, and the lines are not continuous. This is probably influenced by unreliability of assessors' ratings.

Discriminant analysis

For learning phase 1, the first discriminant function is significantly different across groups (chi-square = 22.36, df = 8, p = .004). The discriminant function coefficients indicate that *insufficient performance* is the best predictor for the classification into groups, respectively followed by *mean performance level*, *rate of growth* and *growth*. In total, the discriminant function successfully predicted group membership for 73.1%, see Table 2.

Table 2. *Predicted group membership for learning phase 1(N=26)*

	Low performers	Moderate performers	High performers	Total
Low performers	7	1	2	10
Moderate performers	1	4	1	6
High performers	1	1	8	10

For learning phase 2, the first discriminant function is also significantly different across groups (chi-square = 45.49, df = 8, p = .000). The discriminant function coefficients indicate that *mean performance* is the best predictor for the classification into groups, respectively followed by *insufficient performance*, *growth* and *rate of growth*. In total, the discriminant function successfully predicted group membership for 90.5%, see Table 3.

Table 3. *Predicted group membership for learning phase 2 (N=21)*

	Low performers	Moderate performers	High performers	Total
Low performers	5	1	0	6
Moderate performers	0	6	0	6
High performers	0	1	8	9

Analysis of variance

An analysis of variance (ANOVA) shows that the means of two variables differ significantly ($p < .05$) across the three groups for learning phase 1 as presented in Table 4; Levene's test showed that the variances of the four variables are all homogeneous.

Table 4. *Analysis of variance (ANOVA) for OJT learning phase 1 (N=26)*

	F	df1	df2	Sig.
Mean performance	8.154	2	23	.002
Insufficient performance	.124	2	23	.884
Growth	1.487	2	23	.247
Rate of growth	14.456	2	23	.000

Rather comparable results were found with an analysis of variance (ANOVA) for learning phase 2. Table 5 shows that the means of all variables differ significantly ($p < .05$) across the three groups. Levene's test showed that we may not assume that the variances of *mean performance level* and *insufficient performance* are homogeneous at $p < .05$.

Table 5. *Analysis of variance (ANOVA) for OJT learning phase 2 (N=21)*

	F	df1	df2	Sig.
Mean performance	28.566	2	18	.000
Insufficient performance	3.812	2	18	.042
Growth	7.640	2	18	.004
Rate of growth	8.960	2	18	.002

The results for both learning phases suggest that *mean performance level* and the *rate of growth* are the best predictors for the classification into the groups, although *growth* and *insufficient performance* generally contribute to this classification as well. It should be noticed that the classification into groups are the same for both learning phases. This may have affected the results because some trainees may fit better in two different groups for the two phases.

Differentiation in competence ratings

Next, we investigated to which extent the three groups differ in means on each competence for the two learning phases together. Figure 9 shows that competences that seem to be critical for ATC performance, *safety, efficiency, mental picture, attention management, planning* and *workload management* are very distinctive. These results are confirmed by an analysis of variance (ANOVA): the means of these competences differ significantly ($p < .001$) across the three groups in contrast with competences that seem to be less ATC-related such as *equipment operation, attitude* and *team orientation.*

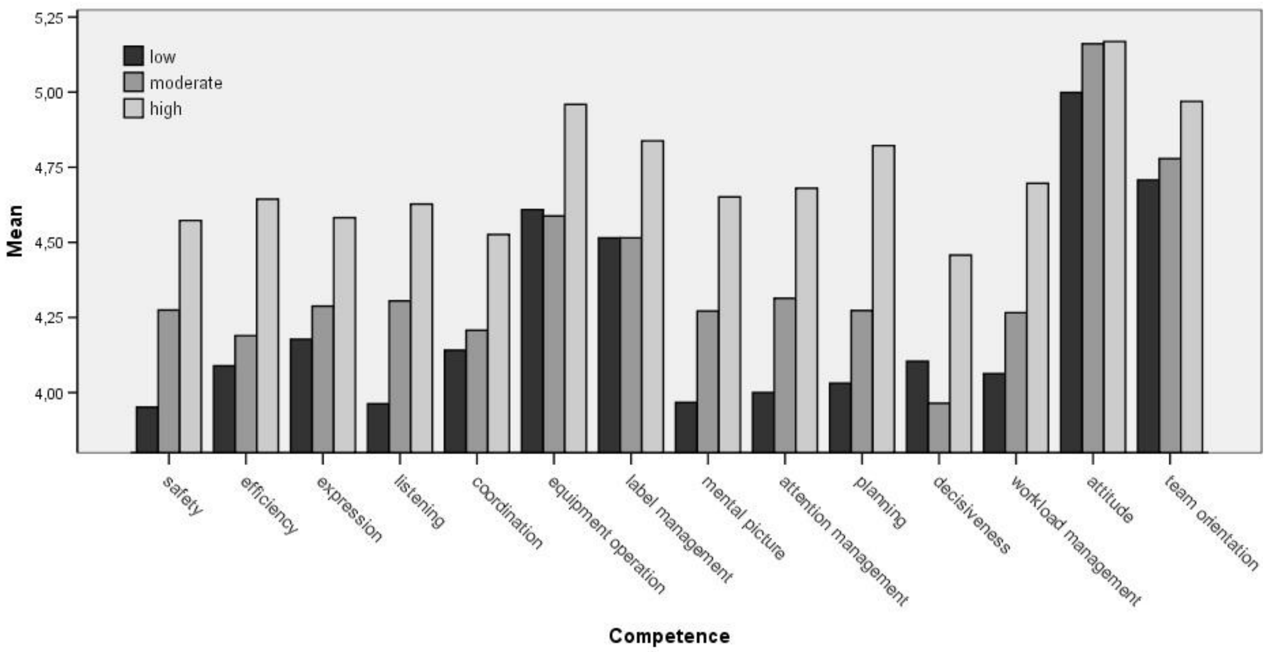

Figure 9. Mean competence ratings (N=27), for two learning phases

Rank order correlations

Finally, we examined competence development for each group of performers by calculating rank order correlations between the variable 'time' and competence ratings. Table 6 generally shows positive correlations. This can be explained by the fact that trainees are assessed against phase level. The low performers who fail in learning phase 2 show progression in learning phase 1 towards the required level, otherwise they would have failed in learning phase 1. The fact that the requirements in learning phase 2 are higher may explain why moderate performers in learning phase 2 show more progression than in learning phase 1. Thus, these findings confirm the results of the learning curve analyses. Differences between competences are not clear; only label management and team orientation do not show much variation.

Table 6. *Rank order correlations (Spearman) between time and competence ratings for learning phase 1 and 2*

Competence	High performers		Moderate performers		Low performers	
	Phase 1 (N = 71)	Phase 2 (N = 81)	Phase 1 (N = 46)	Phase 2 (N = 64)	Phase 1 (N = 85)	Phase 2 (N = 46)
Safety	.12	.23*	-.25	.37**	.21*	-.22
Efficiency	.33**	.28**	-.09	.50**	.27*	.00
Verbal expression	.24*	-.07	-.00	.42**	.18*	.11
Listening	.23*	.15	-.16	.25*	.18*	.13
Co-ordination	.19	.28**	-.20	.27*	.17	.10
Equipment operation	.05	.03	-.06	.30**	.28**	.01
Label management	.24*	.13	.06	.20	.26**	.05
Mental picture	.22*	.26*	.11	.41**	.31**	.02
Attention management	.31**	.39**	.01	.45**	.26**	-.07
Planning	.13	.08	-.11	.30**	.31**	.04
Decisiveness	.12	.38**	-.10	.48**	.22*	.11
Workload management	.28*	.04	.22	.51**	.10	-.04
Attitude	-.16	-31**	-.06	.33**	.18	.08
Team orientation	.03	.07	.03	.08	.10	.16

Discussion and conclusions

Representative learning curves

This chapter proposed a new method to derive learning curves from assessments. We made 'recalibrated' learning curves, based on assessment against performance standards that are constantly recalibrated into the same rating scale. This was needed to examine whether learning curves, produced by the assessment system, were sufficiently representative for learning processes as part of the evaluation. In this way, this study adds to the studies with the Kanfer-Ackerman task, also an ATC task, aimed at modelling learning processes (Ackerman, 1989; Lee & Anderson, 2001; Taatgen & Lee, 2003).

The results have shown that the assessment system is able to represent patterns in learning processes sufficiently. This provides evidence of the quality of the assessment design. We distinguished three groups of trainees based on training success. Their recalibrated learning curves derived from assessment results agree with the three defined patterns of prototypical learning curves in conformance with general learning theory (Newell & Rosenbloom, 1981). Performance usually improves with practice and therefore high and moderate performers show progression over time in the recalibrated learning curves. Low performers achieve a learning plateau earlier than the moderate and high performers. However, the graphs show many highs and lows and they present a certain number of outliers. This confirms that ATC is a very complex skill to acquire (Schneider, 1990), especially in live environments. In this context, we should also take into account that the assessors' ratings are not completely reliable (Oprins, Van Weerdenburg & Burggraaff, 2008; Oprins, 2008). This is very difficult to achieve in OJT, more than in simulator training, due to the ongoing variety of task situations, multiple assessors and many individual differences in learning.

The quantitative analyses with four variables (discriminant analysis, analysis of variance) have shown that classification into the three groups was correctly predicted. The best predictors were *mean performance level* (weighted sum of competence ratings) and *rate of growth* (beta coefficient of the linear regression model). *Mean performance level* is probably a better measure than *insufficient performance* because it indicates how much the trainee performs below or above the standards instead of how often. For progression, *growth* (difference between final and initial performance level) is less distinctive than *rate of growth* because individual learning curves show many variations caused by instable performance of trainees. Rate of growth indicates a certain direction of the learning curve based on the whole range of measures. However, it should be noticed that the beta coefficients are not always reliable since the linear regression model does not fit for each trainee. We also should realize that progression in the recalibrated learning curves differs from that in general learning curves: an horizontal line implies that the trainee is still learning since the required standards increase over time.

Competence development
The three groups of trainees differ much in mean competence ratings. The largest differences were found in competences that are assumed more critical and less trainable such as *mental picture* (cf. Situational Awareness) and *workload management* in comparison with less critical and more trainable competences such as *equipment operation* and *attitude.* This relates to the distinction between respectively 'consistent task components' that depend heavily on individual abilities and 'non-consistent task components' that improve by more practice (Schneider, 1990). ATC is a combination of both. Our findings suggest that some competences are more trainable than others. The three groups also differ in progression on competences. The findings support the analyses of learning curves in a quantitative way.

Practical implications
These findings have some practical implications. The learning curves can be used to adapt training to the trainees' needs, for instance, by recognizing slow starters and (intermediate) learning plateaus. Following progression on singular competences can help to detect specific deficiencies of trainees and to repair them as a next step. More insight should be gained in how to make training adaptive, for instance, by (dynamic) task selection, specific coaching, remedial teaching, re-training, etcetera. Development of self-directed learning skills might help trainees to define what they need by themselves. In addition, pass-fail decisions can be improved by using learning indicators with sufficient predictive validity. Trainees can be classified in one of the three groups based on the characteristics of their (recalibrated) learning curve, and finally predictions can be made about future learning. Ultimately, the measures *mean performance level* and *rate of growth* can be used as a cut-off for pass-fail if predictive validity will have been proven.

General conclusion
In sum, patterns in learning processes can be clearly recognized in the recalibrated learning curves produced by the assessment system CBAS applied in OJT. This implies that CBAS is a well-designed instrument to follow trainees' progression over time, to provide adequate feedback and to develop effective training interventions. The next step is to use the learning curves for improving pass-fail decisions based on quantitative performance measurement. Therefore, research on learning curves and how to derive them from assessment results must be continued, involving a higher number of trainees.

References

Ackerman, P.L. (1989). Individual differences and skill acquisition. In P.L. Ackerman, R.J. Sternberg, & R. Glaser (eds), *Learning and individual differences: advances in theory and research* (pp. 165-217). New York: Freeman and company.

Berk (1986). *Performance assessment: methods and applications.* Baltimore: The Johns Hopkins University.

Lee, F. & Anderson, J.R. (2001). Does learning a complex task have to be complex? A study in learning composition. *Cognitive psychology, 42,* 267-316.

Newell, A., & Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the law of practice. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

O'Connor, P., Hormann, H, Flin, R, Lodge, M., & Goeters, K. (2002). Developing a method for evaluating crew resource management skills: a European perspective. *The international journal of aviation psychology, 12*(3), 263-285.
Oprins, E., Burggraaff, E., & Van Weerdenburg, H. (2006). Design of a competence-based assessment system for air traffic control training. *The international journal of aviation psychology, 16*(3), 297-320.
Oprins, E., Burggraaff, E. & Weerdenburg, H. van (2008). Reliability of assessors' ratings in competence-based air traffic control training. *Human factors and aerospace safety, 6* (4), 305-322.
Oprins, E. (2008). *Design of a competence-based assessment system for air traffic control training.* PhD thesis, Maastricht University.
Schneider, W. (1990). Training high-performance skills: fallacies and guidelines. In M. Venturino (Ed.). *Selected readings in human factors* (pp. 297-311). Santa Monica, CA: The human factors society.
Taatgen, N.A., & Lee, F.L. (2003). Production compilation: a simple mechanism to model complex skill acquisition. *Human factors, 45*(1), 61-76.